

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 19:32:18 ; Search time 88.12 Seconds

(without alignments)  
5572.190 Million cell updates/sec

Title: US-08-961-083-1

Sequence: 1 TAAATCTAGCAGCATATAAA.....ATCCTCAACGACGACAAACA 1999

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NA.\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCITUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1999	100.0	1999	US-08-961-083-1	Sequence 1, Appl
2	1981.4	99.1	2049	US-08-481-435-5	Sequence 5, Appl
3	944.2	47.2	960	US-08-245-511-3	Sequence 3, Appl
4	944.2	47.2	960	US-08-600-993A-3	Sequence 3, Appl
5	117.2	5.9	9100	US-08-743-637B-27	Sequence 27, Appl
6	117.2	5.9	9100	US-08-526-840B-27	Sequence 27, Appl
7	85	4.3	2487	US-08-481-435-1	Sequence 1, Appl
8	62.2	3.1	807	US-08-771-716-1	Sequence 1, Appl
9	62.2	3.1	807	US-08-771-716-3	Sequence 1, Appl
10	62.2	3.1	807	US-09-057-720A-1	Sequence 3, Appl
11	62.2	3.1	807	US-09-057-720A-3	Sequence 3, Appl
12	56.4	2.8	2193	US-08-731-716-1	Sequence 1, Appl
13	56.4	2.8	2193	US-08-731-716-3	Sequence 3, Appl
14	55.2	2.8	428	US-09-060-756-206	Sequence 206, App
15	53.8	2.7	1474	US-09-508-542-17	Sequence 17, Appl
16	51.8	2.6	363	US-09-060-756-626	Sequence 626, Appl
17	49.8	2.5	4403765	US-09-103-840A-2	Sequence 2, Appl
18	49.8	2.5	4411529	US-09-103-840A-1	Sequence 1, Appl
19	45.2	2.3	5511	US-08-928-361B-2	Sequence 2, Appl
20	45.2	2.3	7334	US-08-928-361B-1	Sequence 1, Appl
21	41.8	2.1	5163	US-08-700-651-1	Sequence 1, Appl
22	41.8	2.1	5163	US-08-928-361B-4	Sequence 4, Appl
23	41.8	2.1	5318	US-08-700-651-2	Sequence 2, Appl
24	41.8	2.1	5318	US-08-928-361B-3	Sequence 3, Appl
25	40.6	2.0	223	US-09-060-756-438	Sequence 438, App
26	36.6	1.8	1172	US-08-936-165A-8	Sequence 8, Appl
27	36.6	1.8	7218	US-08-232-463-14	Sequence 14, Appl

C 28	36.4	1.8	473	4	US-09-060-756-671	Sequence 671, App
C 29	35.8	1.8	348	4	US-09-060-756-99	Sequence 99, Appl
C 30	35.2	1.8	43795	3	US-08-742-185-101	Sequence 101, App
C 31	34	1.7	2946	4	US-09-175-928-3	Sequence 3, Appl
C 32	33.8	1.7	1007	4	US-08-858-10668-113	Sequence 113, App
C 33	33.4	1.7	198	5	PCT-US95-10668-3	Sequence 3, Appl
C 34	33.4	1.7	198	5	PCT-US95-10668-4	Sequence 4, Appl
C 35	33.4	1.7	2029	4	US-09-136-574A-46	Sequence 46, Appl
C 36	33.2	1.7	700	4	US-09-236-097-11	Sequence 11, Appl
C 37	33	1.7	700	4	US-09-236-097-8	Sequence 8, Appl
C 38	32.8	1.6	366	4	US-09-060-756-331	Sequence 331, App
C 39	32.8	1.6	1725	1	US-08-257-073-14	Sequence 82, Appl
C 40	32.6	1.6	1543	3	US-08-714-918-82	Sequence 82, Appl
C 41	32.6	1.6	1543	4	US-09-265-315-82	Sequence 82, Appl
C 42	32.6	1.6	1543	4	US-09-265-315-82	Sequence 82, Appl
C 43	32.6	1.6	1543	4	US-09-266-417-82	Sequence 82, Appl
C 44	32.6	1.6	3343	4	US-08-965-762-21	Sequence 21, Appl
C 45	32.4	1.6	3292	1	US-07-814-964-12	Sequence 12, Appl

#### ALIGNMENTS

RESULT 1  
US-08-961-083-1  
Sequence 1, Application US/08961083  
Patent No. 6159469  
GENERAL INFORMATION:  
APPLICANT: Chol et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,083  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1999 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-083-1  
Query Match 100.0%; Score 1999; DB 3; Length 1999;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TAAATCTAGCAGCAATATAAAATCACTGCTGCTGAGTTGGTTGACGCGCGGCA 60  
DB 1 TAAATCTAGCAGCAATATAAAATCACTGCTGCTGAGTTGGTTGACGCGCGGCA 60

QY 61 TGGCCAGCTAATGATATTCOCACAGATTTGGTTAAGGCAATCGTTTCATTCGAAGACCA 120  
| | | | |  
Db 61 TGGCCAGCTAATGATATTCOCACAGATTTGGTTAAGGCAATCGTTTCATTCGAAGACCA 120  
QY 121 TGGCTTCTTCGACACAGGGGGGATTGATACATCCGATCTCTGGAGCTTTCTTGGCCAA 180  
| | | | |  
Db 121 TGGCTTCTTCGACACAGGGGGGATTGATACATCCGATCTCTGGAGCTTTCTTGGCCAA 180  
QY 181 TCGCAAGCAATTCCTCCAGAGTGATGATCACTCTCACCCACAGTTGATTAAGTTGAC 240  
| | | | |  
Db 181 TCGCAAGCAATTCCTCCAGAGTGATGATCACTCTCACCCACAGTTGATTAAGTTGAC 240  
QY 241 TTAATTTTCACTTCGACCTTCGACACGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300  
| | | | |  
Db 241 TTAATTTTCACTTCGACCTTCGACACGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300  
QY 301 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAAATCTTGACTACTATTAATAA 360  
| | | | |  
Db 301 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAAATCTTGACTACTATTAATAA 360  
QY 361 GGTCTCATGTCTAATGGGAACTATGAAATGACAGAGGCTCAAAAGCTACTATGTTAA 420  
| | | | |  
Db 361 GGTCTCATGTCTAATGGGAACTATGAAATGACAGAGGCTCAAAAGCTACTATGTTAA 420  
QY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGGCTTGTGGTGAATGGCTCAGGACCC 480  
| | | | |  
Db 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGGCTTGTGGTGAATGGCTCAGGACCC 480  
QY 481 AAACCAATATGACCCCTATTCACATCCAGAGAGAGCCCAAGAGCCGCAAACTTGGCTT 540  
| | | | |  
Db 481 AAACCAATATGACCCCTATTCACATCCAGAGAGAGCCCAAGAGCCGCAAACTTGGCTT 540  
QY 541 ATCTGAATATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTAATAC 600  
| | | | |  
Db 541 ATCTGAATATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTAATAC 600  
QY 601 ACCAATTAATGATGACCTCAAAAGTCTCAATTCAGCAAGTAAATTCCTGCTTACATGA 660  
| | | | |  
Db 601 ACCAATTAATGATGACCTCAAAAGTCTCAATTCAGCAAGTAAATTCCTGCTTACATGA 660  
QY 661 TTAATTAACCTCAAGAGATGATCAATCAAGTGAAGAAAGCAAGGCTATTAACCTACTCAC 720  
| | | | |  
Db 661 TTAATTAACCTCAAGAGATGATCAATCAAGTGAAGAAAGCAAGGCTATTAACCTACTCAC 720  
QY 721 AACTGGAGATGATGCTTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780  
| | | | |  
Db 721 AACTGGAGATGATGCTTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780  
QY 781 TTACATATACAGAGATGATGCTTACCAAGATGATGATGATGATGATGATGATGATGAT 840  
| | | | |  
Db 781 TTACATATACAGAGATGATGCTTACCAAGATGATGATGATGATGATGATGATGATGAT 840  
QY 841 TGTGTATGTTTCTAAGGATTAAGTCTTGTCCAGCTAGAGACAGCCGATCAATCAAGTAA 900  
| | | | |  
Db 841 TGTGTATGTTTCTAAGGATTAAGTCTTGTCCAGCTAGAGACAGCCGATCAATCAAGTAA 900  
QY 901 TGTTCCTCTCGGAATTAACCAAGAGTAAACCAAGAGTGGGATGCAATCAATGAA 960  
| | | | |  
Db 901 TGTTCCTCTCGGAATTAACCAAGAGTAAACCAAGAGTGGGATGCAATCAATGAA 960  
QY 961 ACCGATACAGAGTATGCTGCTGCTTGGAGTACGCTGCTACGATTCAGTCTACTAT 1020  
| | | | |  
Db 961 ACCGATACAGAGTATGCTGCTGCTTGGAGTACGCTGCTACGATTCAGTCTACTAT 1020  
QY 1021 CGTTACAGATGAGCCCTATTAACCTTGGGACAAATTAATCTCTGTTATTAATCGGATAG 1080  
| | | | |  
Db 1021 CGTTACAGATGAGCCCTATTAACCTTGGGACAAATTAATCTCTGTTATTAATCGGATAG 1080  
QY 1081 GGGCTACTTGGCAACATACCTTGCATATACGCTGCAACAATGCGGAAACGCTCCACG 1140  
| | | | |  
Db 1081 GGGCTACTTGGCAACATACCTTGCATATACGCTGCAACAATGCGGAAACGCTCCACG 1140

QY 1141 CGTGAAGAACTTAACAGAGCTGAGACTCAACCGCGCCAGACCTTCTTAATAGTCTAGG 1200  
| | | | |  
Db 1141 CGTGAAGAACTTAACAGAGCTGAGACTCAACCGCGCCAGACCTTCTTAATAGTCTAGG 1200  
QY 1201 AATGCACTACCCCAAGTATTCACCTACCTCAATGGCATTTCGAATGACCAACCGAATCAGA 1260  
| | | | |  
Db 1201 AATGCACTACCCCAAGTATTCACCTACCTCAATGGCATTTCGAATGACCAACCGAATCAGA 1260  
QY 1261 CAAAAATATGAGCAAGTATGAAAGAGTGGCTGCTTACGCTTGCCTTTCGCAATGG 1320  
| | | | |  
Db 1261 CAAAAATATGAGCAAGTATGAAAGAGTGGCTGCTTACGCTTGCCTTTCGCAATGG 1320  
QY 1321 TGGAACTTACTATTAACCAATGATATATCAATTAAGTCTTAAAGTATGATGAGTGAATA 1380  
| | | | |  
Db 1321 TGGAACTTACTATTAACCAATGATATATCAATTAAGTCTTAAAGTATGATGAGTGAATA 1380  
QY 1381 AGAGTTCTCTAATGTGGAACTGCTGCATGAAAGGAAAGCAGACAGCTATATATGACCGA 1440  
| | | | |  
Db 1381 AGAGTTCTCTAATGTGGAACTGCTGCATGAAAGGAAAGCAGACAGCTATATATGACCGA 1440  
QY 1441 CATGATGAAAAAGCTCTGACTATGTAAGTGAAGTGAAGAAATGCTATCTTGTGGCTGCC 1500  
| | | | |  
Db 1441 CATGATGAAAAAGCTCTGACTATGTAAGTGAAGTGAAGAAATGCTATCTTGTGGCTGCC 1500  
QY 1501 TCAGGCTGTTAAACAGGAACCTCTAATATATACAGACGAGAAATTTGAAAAACCATATCAA 1560  
| | | | |  
Db 1501 TCAGGCTGTTAAACAGGAACCTCTAATATATACAGACGAGAAATTTGAAAAACCATATCAA 1560  
QY 1561 GACCTCTCAATTTGTAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
| | | | |  
Db 1561 GACCTCTCAATTTGTAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
QY 1621 GCGTGTATGAGACAGGCTATTTCTAACCCTGACACCACTGTATGAGCAATGGCTTACGAT 1680  
| | | | |  
Db 1621 GCGTGTATGAGACAGGCTATTTCTAACCCTGACACCACTGTATGAGCAATGGCTTACGAT 1680  
QY 1681 CGCTGCAAAAGTTTACCGCTATATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
| | | | |  
Db 1681 CGCTGCAAAAGTTTACCGCTATATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
QY 1741 TTGGAAATATACAGAGAGGCTCTACAGAAATGGAAGAAATTCGATTTAAAAATGCTCTCG 1800  
| | | | |  
Db 1741 TTGGAAATATACAGAGAGGCTCTACAGAAATGGAAGAAATTCGATTTAAAAATGCTCTCG 1800  
QY 1801 TTCTACGTGAAGTCACTGCTGCTCCACAAACCCCATCACTGAAGTTCAGATCATC 1860  
| | | | |  
Db 1801 TTCTACGTGAAGTCACTGCTGCTCCACAAACCCCATCACTGAAGTTCAGATCATC 1860  
QY 1861 ATCAGATAGTTCAACTTCAAGTCTAGCTCAACCACTCCAAAGCAAAATTAATAGTACGAC 1920  
| | | | |  
Db 1861 ATCAGATAGTTCAACTTCAAGTCTAGCTCAACCACTCCAAAGCAAAATTAATAGTACGAC 1920  
QY 1921 TACCAATCTTAACAAATTAATGCAACAAATCAATTAACAAATCAATTAATGAA 1980  
| | | | |  
Db 1921 TACCAATCTTAACAAATTAATGCAACAAATCAATTAACAAATCAATTAATGAA 1980  
QY 1981 TCCTCAACGACGACCA 1999  
| | | | |  
Db 1981 TCCTCAACGACGACCA 1999

RESULT 2  
US-08-481-435-5  
: Sequence 5. Application us/08481435  
: Patent No. 6027906  
: GENERAL INFORMATION:  
: APPLICANT: Balgatesh, Tanjore S  
: APPLICANT: Town, Christine  
: TITLE OF INVENTION: No. 6027906el Polypeptides  
: NUMBER OF SEQUENCES: 42  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: White & Case  
: STREET: 1155 Avenue of the Americas

```
? CITY: New York
? STATE: New York
? COUNTRY: United States
? ZIP: 10036-2787
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/481,435
? FILING DATE: 10-JUL-1995
? CLASSIFICATION: 435
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: IN 580/MAS/94
? FILING DATE: 01-JUL-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: SE 9404072-2
? FILING DATE: 24-NOV-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Sterner, Richard J.
? REGISTRATION NUMBER: 35,372
? REFERENCE/DOCKET NUMBER: 1103326-151
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 819-8783
? TELEFAX: (212) 354-8113
? INFORMATION FOR SEQ. ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2049 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? ORIGINAL SOURCE:
? ORGANISM: Streptococcus pneumoniae
? STRAIN: PM 1
? IMMEDIATE SOURCE:
? LIBRARY: PCR cloning
? CLONE: PARC 0512 Soluble PBP 1A del 38
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..2049
? FEATURE:
? NAME/KEY: mat.peptide
? LOCATION: 1..2046
? US-08-481-435-5

Query Match          99.1%; Score 1981.4; DB 3; Length 2049;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

Db 348 AGCGATTCAAGTGAACAAAGCAACCAAGCAAGAAATCTGTGACCTACTATATAATAA 407
QY 361 GGTCACATGTCTATATGGGAACCTATGGAATGCAGACAGCGCTCAAACTACTATGTAA 420
Db 408 GGTCACATGTCTATATGGGAACCTATGGAATGCAGACAGCGCTCAAACTACTATGTAA 467
QY 421 AGACCTCAATTAATTAAGTTTACCTGCTGTAAGCTTCTGCTGGAATGCTCAGGCACC 480
Db 468 AGACCTCAATTAATTAAGTTTACCTGCTGTAAGCTTCTGCTGGAATGCTCAGGCACC 527
QY 481 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGCCGCGAACTTGCTT 540
Db 528 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGCCGCGAACTTGCTT 587
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTAATAC 600
Db 588 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTAATAC 647
QY 601 ACCAATTACGTAGGACTACAAAGCTCAATGACGAAGTAATTAACCTGCTTACATGA 660
Db 648 ACCAATTACGTAGGACTACAAAGCTCAATGACGAAGTAATTAACCTGCTTACATGA 707
QY 661 TAAATTACCTGAAGAGTCAATCAATCAAGTTGAAGAAAGAGGCTAATACCTACTAC 720
Db 708 TAATTACCTGAAGAGTCAATCAATCAAGTTGAAGAAAGAGGCTAATACCTACTAC 767
QY 721 AACTGGAGTATGCTTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
Db 768 AACTGGAGTATGCTTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 827
QY 781 TTACAATACAGAGAAATACGTTGCTATCCAGAGAAATGCAATGCAATGCTTCTACAT 840
Db 828 TTACAATACAGAGAAATACGTTGCTATCCAGAGAAATGCAATGCAATGCTTCTACAT 887
QY 841 TGTGATGTTTCTTAACGCTTAAAGCTATGCTGCTGAGAGACAGCCATCAGTCAAGTAA 900
Db 888 TGTGATGTTTCTTAACGCTTAAAGCTATGCTGCTGAGAGACAGCCATCAGTCAAGTAA 947
QY 901 TGTTCCTTGGGAATTAACCAAGCAGTAGAAGCAAAACCGGAGTGGGATCAACTATGAA 960
Db 948 TGTTCCTTGGGAATTAACCAAGCAGTAGAAGCAAAACCGGAGTGGGATCAACTATGAA 1007
QY 961 ACCGATCAGACGACTATGCTCTGCTGAGTACGGTGTACGAGTACGCACTAT 1020
Db 1008 ACCGATCAGACGACTATGCTCTGCTGAGTACGGTGTACGAGTACGCACTAT 1067
QY 1021 CGTTACAGATGAGCCCTATTAACCTGCGGACAAATACCTGTTATTAACCTGGATAG 1080
Db 1068 CGTTACAGATGAGCCCTATTAACCTGCGGACAAATACCTGTTATTAACCTGGATAG 1127
QY 1081 GGGCTACTTTGGCAACATCACTTGCAATAGCCCTGCAACAAATCGGAAACGTCCAGC 1140
Db 1128 GGGCTACTTTGGCAACATCACTTGCAATAGCCCTGCAACAAATCGGAAACGTCCAGC 1187
QY 1141 CGTGAACCTCTAAACAGGTCGAGTCAACCGCGCCCAACACTTTCTTAATGCTCAGG 1200
Db 1188 CGTGAACCTCTAAACAGGTCGAGTCAACCGCGCCCAACACTTTCTTAATGCTCAGG 1247
QY 1201 AATGACCTACCAAGTATTAATCTACTACTCAATATGCAATTCCTTAAGTAACCAACCAATCAGA 1260
Db 1248 AATGACCTACCAAGTATTAATCTACTACTCAATATGCAATTCCTTAAGTAACCAACCAATCAGA 1307
QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAAAGATGCTGCTTACAGCTGCTCTTGCATAATGG 1320
Db 1308 CAAAAAATATGAGCAAGTAGTGAAGAAAGATGCTGCTTACAGCTGCTCTTGCATAATGG 1367
QY 1321 TGAACCTTACTATTAACCAATGTAATATCCATTAAGTGTGCTTAAGTATGAGTGAAGTAA 1380
Db 1368 TGAACCTTACTATTAACCAATGTAATATCCATTAAGTGTGCTTAAGTATGAGTGAAGTAA 1427
QY 1381 AGAGTCTCTAATGTCGGAAGCTGCTGCAAGGAAGCAACAGCCATATGATGACCGA 1440
```





|||||  
Db 612 ACCAATTACGATGGGCTACAAAGTCTCAATATAGCAGATTAATACCTCTTACATGGA 671  
QY 661 TAAATACCTCAAGGAAGTCATCAATCAAGTTGAAGAAAGAAAGAGCTATTAACCTACTAC 720  
Db 672 TAAATACCTCAAGGAAGTCATCAATCAAGTTGAAGAAAGAAAGAGCTATTAACCTACTAC 731  
QY 721 AACTGGAGATGATGCTACCAATGTAGACCAAGAAAGCTCAAAACATCTGTGGGATAT 780  
Db 732 AACTGGAGATGATGCTACCAATGTAGACCAAGAAAGCTCAAAACATCTGTGGGATAT 791  
QY 781 TTACATATACAGCAATACGTTGCTATCCAGACGATGAATGCAAGTCCCTTACCAT 840  
Db 792 TTACATATACAGCAATACGTTGCTATCCAGACGATGAATGCAAGTCCCTTACCAT 851  
QY 841 TGTGTATGTTTCTAACGATTAAGTCATTTGCCAGCTAGAGACGCCATCAGTCAAGTAA 900  
Db 852 TGTGTATGTTTCTAACGATTAAGTCATTTGCCAGCTAGAGACGCCATCAGTCAAGTAA 911  
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTGAAGAAACCCGGAGCTGGGGA 949  
Db 912 TGTTCCTTCGGAATTAACCAAGCAGTGAAGAAACCCGGAGCTGGGGA 960

RESULT 4  
US-08-600-993A-3  
; Sequence 3, Application US/08600993A  
; Patent No. 5981229  
; GENERAL INFORMATION:  
; APPLICANT: Masure, H Robert  
; APPLICANT: Pearce, Barbara J  
; APPLICANT: Tuomonen, Elaine  
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND  
; NUMBER OF SEQUENCES: 59  
; ADDRESS: Klauher & Jackson  
; ADDRESS: Klauher & Jackson  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/600,993A  
; FILING DATE: 1-MAR-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/245,511  
; FILING DATE: 18-MAY-1994  
; CLASSIFICATION: 435  
; APPLICATION DATA:  
; APPLICATION NUMBER: US 08/116,541  
; FILING DATE: 01-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-069 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 960 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE: Streptococcus pneumoniae  
; STRAIN: R6  
; IMMEDIATE SOURCE: SPUR42  
; CLONE: SPUR42  
; FEATURE: CDS  
; NAME/KEY: CDS  
; LOCATION: 1..960  
; US-08-600-993A-3

Query Match 47.2%; Score 944.2; DB 2; Length 960;  
Best Local Similarity 99.7%; Pred. No. 1.9e-278;  
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATTAATCAATCAATGCTGATGGTGGTTGGAACGGCGGCA 60  
Db 12 TAAATCTACGACAAATTAATCAATCAATGCTGATGGTGGTTGGAACGGCGGCTCA 71  
QY 61 TGCCCAAGCTAATGATATTCACAGATTTGTTAAGCAATGCTTCTATCGAAGCA 120  
Db 72 TGCCCAAGCTAATGATATTCACAGATTTGTTAAGCAATGCTTCTATCGAAGCA 131  
QY 121 TCGCTTCTTGACACAGGGGGATTTGATACCATCCGATCTTGGAGCTTTTGGCA 180  
Db 132 TCGCTTCTTGACACAGGGGGATTTGATACCATCCGATCTTGGAGCTTTTGGCA 191  
QY 181 TCGCAAGCAATTCCTCCAGGTGGATCACTCCACCAAGTGTATTAAGTTGAC 240  
Db 192 TCGCAAGCAATTCCTCCAGGTGGATCACTCCACCAAGTGTATTAAGTTGAC 251  
QY 241 TTACTTTCACTTCGACTTCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGT 300  
Db 252 TTACTTTCACTTCGACTTCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGT 311  
QY 301 AGCGATTCACTTGAACAAAGCAACCAAGCAAGAAATCTTCACTATTAATAA 360  
Db 312 AGCGATTCACTTGAACAAAGCAACCAAGCAAGAAATCTTCACTATTAATAA 371  
QY 361 GGTCTACATGCTTAATGGGAGTATGGAATGACAGACAGCTCAAACTATGTTAA 420  
Db 372 GGTCTACATGCTTAATGGGAGTATGGAATGACAGACAGCTCAAACTATGTTAA 431  
QY 421 AGACTCAATATTTAAGTTTACCTCAGTTAGGCTTGGTGGTGAATGCTCAGGAC 480  
Db 432 AGACTCAATATTTAAGTTTACCTCAGTTAGGCTTGGTGGTGAATGCTCAGGAC 491  
QY 481 AAACCAATATGACCCCTATTACATCCAGAGCAAGCCCAAGACCCGGAATTTG 540  
Db 492 AAACCAATATGACCCCTATTACATCCAGAGCAAGCCCAAGACCCGGAATTTG 551  
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGTGAACATATGAAGAAACCAATAC 600  
Db 552 ATCTGAATGAAAAATCAAGGCTACATCTCTGTGAACATATGAAGAAACCAATAC 611  
QY 601 ACCAATTACGATGAGTCAAGTCAATAGCAAGTCAATTAACCTCTTACATGGA 660  
Db 612 ACCAATTACGATGAGTCAAGTCAATAGCAAGTCAATTAACCTCTTACATGGA 671  
QY 661 TAAATACCTCAAGGAAGTCATCAATCAAGTTGAAGAAAGAAAGAGCTATTAAC 720  
Db 672 TAAATACCTCAAGGAAGTCATCAATCAAGTTGAAGAAAGAAAGAGCTATTAAC 731  
QY 721 AACTGGAGATGATGCTACCAATGTAGACCAAGAAAGCTCAAAACATCTGTGG 780  
Db 732 AACTGGAGATGATGCTACCAATGTAGACCAAGAAAGCTCAAAACATCTGTGG 791  
QY 781 TTACATATACAGCAATACGTTGCTATCCAGACGATGAATGCAAGTCCCTTAC 840  
Db 792 TTACATATACAGCAATACGTTGCTATCCAGACGATGAATGCAAGTCCCTTAC 851





Db 200 GCTTCTACGACATCAGCGCGCTTGACCCGGTGGGGATCTCCGTGACAGCAAGCGTGCGC 259  
QY 183 TGAAGCAATTCCTCCAGAGTGATGATCACTTCACCCAGATGTAATGATGACTT 242  
Db 260 TGTTCCTCGGTCACGCGTCACAAAGGGCAAGTACCATTAACCCAGAGCTGGCGAGAACT 319  
QY 243 ACTTTCAACTTCGACTTCGACAGCAAGTATTTCTGTAAGGCTCAGAAAGCTTGTTAG 302  
Db 320 TCTTCCTCAGTCC-----AGAACGACGCTGATGCCGTAAAGATTAAAGAACTCTTCG 373  
QY 303 CGATTCACTTGAACAAAAGCAACCAAGCAAGAAATCTTACCTACTATTAATAAG 362  
Db 374 CGATTCCATTGAACAGCTGCTGACGAAAGAGATCCTGAGAGTTATCTGAAACAAGA 433  
QY 363 TCTCATGTCTAATGGAGACTATGTAATGCAGACAGCAGCTCAAAACTACTAGTGTAAAG 422  
Db 434 TTTACCTTGTTACCGCGCTTATGTTGCTGCTGCGGCAAGCTTAATTTCCGAAATA 493  
QY 423 ACCTCAATTAATTTAATTTTACCTGATTAAGCTTGTGCTGGAATGCTCAGGACCA 482  
Db 494 CGGTGACCACTGACGCTGAAAGAAATGGGGGTATGATAGCGGGCTGCCGAAAGCGCTT 553  
QY 483 ACCAATATGACCCCTATTACATCCAGAGCAAGCCCAAGCCGGAATTTGGTCTTAT 542  
Db 554 CCACCTTCAACCCCTCTACTGATGATCGTCCGCTGCGCGGTAACTGCTGTGT 613  
QY 543 CTGAATGAATAATCAAGGCTACATCTGCTGAACAGTATGAAAGACACTCAATAC 602  
Db 614 CGCGAATGCTGATGAAAGGTAATACCCCAACAGTGTGATCAGACAGCACTGAGG 673  
QY 603 CAATTAAGTAT 613  
Db 674 CGATTAACGCT 684

RESULT 8  
US-08-771-716-1  
Sequence 1, Application US/08771716  
Patent No. 5922540  
GENERAL INFORMATION:  
APPLICANT: Peery, Robert B.  
APPLICANT: Hoskins, Joann  
APPLICANT: Jaskunas, S. Richard  
APPLICANT: Skatrud, Paul L.  
TITLE OF INVENTION: Monofunctional Glycosyltransferase  
TITLE OF INVENTION: Gene of Staphylococcus Aureus  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/771.716  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11067  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 807 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..807  
US-08-771-716-1

Query Match 3.1%; Score 62.2; DB 2; Length 807;  
Best Local Similarity 49.9%; Pred. No. 3.4e-09;  
Matches 191; Conservative 0; Mismatches 183; Indels 9; Gaps 1;

QY 67 AGCTAATATATTTCCACAGATTTGGTTAAGGCATCTTTCTATACGAAGCAATCGCTT 126  
Db 252 AGCTGATTAACATGCCAGAGTATGTTAAAGTGCTTATTTCAATGAAGAATGAACGATT 311  
QY 127 CTTCGACCACAGGGGATGATACCATCCGATATCCTGGAGCTTCTTCCGCAATCTCA 186  
Db 312 CTACATATATGATGATTCGATTTGAAGGTACAACTAGAGCTTATTTTCAACGATTAG 371  
QY 187 AAGCAATTCCTCCAGGTGATCACTCTCACCCACAGTTGATTAAGTACTTACTT 246  
Db 372 CGACAGATGTCGAAGGTGATGATCAATTAACACAAAGTTGCAAAATTAATTTTA 431  
QY 247 TTCAACTTCGACTTCGACACAGACTATTTCTGTAAGCTCAGAGACTTGGTTAGCAT 306  
Db 432 TGATTAAT-----GATCCTTCAATTTAGTAAGAAAGTAAAGATTAATTTGCTCA 482  
QY 307 TCAGTTAGACAAAAGCAACCAAGCAAGAAATCTTCACTACTATTAATAAGTCTA 366  
Db 483 TCGAGTTTAAACATATATTAAGAAAGAAATTTAAGCTTTTAAATTAATTAATTTA 542  
QY 367 CATGCTAATGGGAATGGAATGCAAGACAGACAGCTCAAAACTACTATTTGTTAAAGACT 426  
Db 543 CTTTGGGATTAATCAATTAATGCTTTAGCGGCGGACGCAAAACCATTAATTTGGAACACCGT 602  
QY 427 CAATTAATTAATTTTACTCTAGT 449  
Db 603 GAATTAATAATGATCAACCAATGT 625

RESULT 9  
US-08-771-716-3  
Sequence 3, Application US/08771716  
Patent No. 5922540  
GENERAL INFORMATION:  
APPLICANT: Peery, Robert B.  
APPLICANT: Hoskins, Joann  
APPLICANT: Jaskunas, S. Richard  
APPLICANT: Skatrud, Paul L.  
TITLE OF INVENTION: Monofunctional Glycosyltransferase  
TITLE OF INVENTION: Gene of Staphylococcus Aureus  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/771.716  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:

```

; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39, 872
; REFERENCE/DOCKET NUMBER: X-11067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-771-716-3

```

```

Query Match          3.1%; Score 62.2; DB 2; Length 807;
Best Local Similarity 33.9%; Pred. No. 3.4e-09;
Matches 130; Conservative 61; Mismatches 183; Indels 9; Gaps 1;

```

```

QY 67 AGCTATGATATTCCTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCATGCTT 126
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 252 AGCUGAUAACAUCCAGAGUAGUUAAGGUGCCUUUAUUCAUUGAAGUAGACGAAU 311
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 CTTCGACACAGGGGATTCATCCATCCGCTATCTGGAGCTTCTTGGCAATCTGCA 186
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 312 CUCACAAUCCAUCAUGAUUCCAUUAGGUAACACUAGGCUUUUAUUCAACGAAUAG 371
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 AAGCAATTCCTCCAGAGGATGATCACTCTACCCCAACAGTTGATTAAGTTGACTTACT 246
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 372 CGACAGAGAUUGGCAAGUGGUGAUCACAUACAAAGUUGUCAAUAUUUAUUUA 431
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 247 TTCACATTCGACTCCGACGACCTATTTCTCTAAGGCTCAGAGAGCTTGGTTAGCGAT 306
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 432 UGAUAU-----GAGCGUUCAUUUUACUAGAAAGUAAAGUUAUUUUGUAGCUA 482
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 307 TCAGTTAGAACAAAGCAACCAAGCAAGAACTTGACCTACTATATATAATAGACTA 366
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 483 UCGAGUUGAAACAAUUAUUAUUAAGCAAAUUAUUAAGCUUUUAUUUAUUAUUA 542
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 367 CATGCTAATGGGAGACTGATGATGACAGACGAGCTCAAAACTACTATGTTAAGACT 426
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 543 CUUUGGGAUAUACAUUAUUGCGUGAGGCGGACGAAACCAUUAUUUGAACAACCGU 602
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 427 CAATAATTAAATTACTTCAGT 449
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 603 GAUAAAAAUAGUACAACAAGU 625
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 10
US-09-057-720A-1
; Sequence 1, Application US/09057720A
; Patent No. 6143868
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Hoskins, Joann
; APPLICANT: Jaskunas, S. Richard
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: Monofunctional Glycosyltransferase Gene
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/057,720A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..807
; US-09-057-720A-1

```

```

Query Match          3.1%; Score 62.2; DB 3; Length 807;
Best Local Similarity 49.9%; Pred. No. 3.4e-09;
Matches 191; Conservative 0; Mismatches 183; Indels 9; Gaps 1;

```

```

QY 67 AGCTATGATATTCCTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCATGCTT 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 252 AGCTGATATGATGCGAGATATGTTAAGGTCCTTATTTTCAATGGAAGATGACATTT 311
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 127 CTTCGACACAGGGGATGATACATCCGATTCCTGGAGCTTCTTGGCAATCTGCA 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 312 CTACAAATCATCATGATTTGATTTGAAGGTAACAAGCTTTATTTCAACGATTAAG 371
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 187 AAGCAATTCCTCCAGAGGATGATCACTCTACCCCAACAGTTGATTAAGTTGACTTACT 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 372 CGACAGAGATGTGCAAGGTGTGTACCATTTACCAACAGTTGTCAAAATATTTT 431
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 247 TTCACATTCGACTCCGACGACCTATTTCTGTAAGGCTCAGAGAGCTTGGTTAGCGAT 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 432 TGATAT-----GATGCTTCATTTACTAGAAAAGTAAAGATTTTGTAGCTCA 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 307 TCAGTTAGAACAAAGCAACCAAGCAAGAACTTGACCTACTATATAATAAGGCTTA 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 483 TCAGTTGAAACAAATATATAAGAACGAATTTTAAGCTTTATTAATATATTTTA 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 367 CATGCTAATGGGAGACTGATGATGACAGACGAGCTCAAAACTACTATGTTAAGACT 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 543 CTTTGGGATATCAATATACGCTTGAGGGCGGACAAACCATTAATTTGAAACACGT 602
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 427 CAATAATTAAATTACTTCAGT 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 603 GAATAAAATGTAACAATGT 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 11
US-09-057-720A-3
; Sequence 3, Application US/09057720A
; Patent No. 6143868
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Hoskins, Joann
; APPLICANT: Jaskunas, S. Richard
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: Monofunctional Glycosyltransferase Gene
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company

```



Query Match	Best Local Similarity	Score	DB 1	Length	2193
Matches 270; Conservative 111; Mismatches 446; Indels 27; Gaps	31.6%;	56.4;	DB 1;	2193;	
Query	5	ATCTAGCACAATTAATAATCACTCATCTGCTGACTGCTGGTTCTGAAACCGCGCTCAATGCC	64		
Db	283	AUUUUUAGCGGUGAGAAAAAGAGCGUGGUCUUUUUAGGUAAGGAAACGUAUUUU	342		
Qy	65	CAACCTATATATATCTCCACAGATTTGGTTGAAGCAATGTTCTTATGCAACGATGCG	124		
Db	343	GAGCGACGACAGCAACAGUAAAAAUCUGGCACAAAGCUGUUUUUUGCGACAGAAACGCUUCU	402		
Qy	125	TTCTTCACACAGAGGGGATTGATTAACATCCGATCTGGAGCTTTCTTGGCAATCTG	184		
Db	403	UUCUAUUAAAAUAGCGGUAUUAAACUAUGGCGGUUC-----UUCUGGCUAUGUC	453		
Qy	185	CAAAAGCAATTCCTCCAAAGGTGATCACTCTCACCAACAGTTGATTAAGTTGACTTAC	244		
Db	454	ACUGCGGACGUGUACGUGUGGUCUCCACUAUACCAACAGCUGCUAAAAACGCUAU	513		
Qy	245	TTTTCAACTTCGACTTCGACACGACTATTTCTGTAAGGCTCAGCAACTTGTGTAGCG	304		
Db	514	UUAA-----CGCAGGAUCAAACGUGUAGAGAAAAAGCGAAAGAAUUUUUCCUUGCC	564		
Qy	305	ATTGAGTTAGCAAAAAGCAACCAAGCAAAATCTGTGACCTACTATTAATAAGATC	364		
Db	565	UUGAUAUUAAAGCAAAAUAUAGUAAAGAACCAAUUUAACCAUGUACCUUAACACGCU	624		
Qy	365	TACATGCTCATGTGGAATGGAATGACAGACAGCTCAAAAACACTATGATGTAAGAC	424		
Db	625	UUUUUUGGAUUUGUGUGGGUGUAGAAUGACGACGUAAGAAUAUUUGUGAUGUUCU	684		
Qy	425	CTCATATATTTAAAGTTTACTCTGACTTACGCTTCTGCTGCTGGAATGCTTACGACCAAC	484		
Db	685	GCAACAGAAAGUGAGUUGUAGUAAAGCUGCGACUCUGGACGAGGAGUGGCUAAAGGCGGAA	744		
Qy	485	CAATATGACCCCTATTTACATTCACAGAAAGAGCCCAAGACCGCGAATCTGGCTTATCT	544		
Db	745	CUGAUAUUCCCUUAGAUUCCGUAAGAUUUCACUAUUUGGGCGAUUCCUGUCAG	804		
Qy	545	GAATGAAAAATCAAGGCTCATCTGCTGCAACAGTATGAGAAAGCACTCAATACCA	604		
Db	805	AUUUUUGUUCGACGAGGAUUUUUAGUAAAAAACCAAGAAACGAAUCUGUAGUUGAU	864		
Qy	605	ATTACTGATGACTACCAAGTCTCAATCAGCAAGTATTAATTAATTAATTAATTAAT	664		
Db	865	AUGACUUCGCAUUUGCAGUAUGAUUAGAAAGAAAUUCUACAUUACCGUUAACCCUUCU	924		
Qy	665	TACTCTCAAGGAAGTCATCAATGTTG-----AAGAGAAACAGGCTATTAACCTA	715		
Db	925	UUUUUUUAGUCGGUGGUUUUAGAGCUGUUUCCAAUUAUUUUUACAGAGGAAGAGAU	984		
Qy	716	CTCAACACTGGGATGATGTCCTACACAATGTAGACCAAGAAAGCTCAAAACATCTGTG	775		
Db	985	GUCUAUUUAGGCUAGCCGAUUUUACACAGAGCUGGACCAAAACUACCAAGCAAAUUAUG	1044		
Qy	776	GATATTTTACATACAGACGAATACGTTGCCTATCAGACAGATGAAATTTGCAAGTCCGTTCT	835		
Db	1045	AUUUUUUUAGAAAAACAGAUUGCUUUUCCGAGGCGAGAGAUAGAAACGUUUGCUACUA	1104		
Qy	836	ACCAATTTGTTGATGT 849			
Db	1105	GGAAGUGUAGUCUU 1118			









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 11, 2002, 01:34:34 ; Search time 3439.02 Seconds  
(without alignments)  
12575.723 Million cell updates/sec

Title: US-08-961-083-1  
Perfect score: 1999  
Sequence: 1 TAAATCTGACGACATATAA.....ATTCCTACACGACACACCA 1999

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/2/pna/PCRTUS.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/pna/US06.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/pna/US07.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/pna/US081.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/pna/US081.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/pna/US082.COMB.seq:\*  
7: /cgn2\_6/ptodata/2/pna/US083.COMB.seq:\*  
8: /cgn2\_6/ptodata/2/pna/US084.COMB.seq:\*  
9: /cgn2\_6/ptodata/2/pna/US085.COMB.seq:\*  
10: /cgn2\_6/ptodata/2/pna/US086.COMB.seq:\*  
11: /cgn2\_6/ptodata/2/pna/US087.COMB.seq:\*  
12: /cgn2\_6/ptodata/2/pna/US088.COMB.seq:\*  
13: /cgn2\_6/ptodata/2/pna/US089.COMB.seq:\*  
14: /cgn2\_6/ptodata/2/pna/US090.COMB.seq:\*  
15: /cgn2\_6/ptodata/2/pna/US091.COMB.seq:\*  
16: /cgn2\_6/ptodata/2/pna/US092.COMB.seq:\*  
17: /cgn2\_6/ptodata/2/pna/US093.COMB.seq:\*  
18: /cgn2\_6/ptodata/2/pna/US094.COMB.seq:\*  
19: /cgn2\_6/ptodata/2/pna/US095.COMB.seq:\*  
20: /cgn2\_6/ptodata/2/pna/US095B.COMB.seq:\*  
21: /cgn2\_6/ptodata/2/pna/US095C.COMB.seq:\*  
22: /cgn2\_6/ptodata/2/pna/US095D.COMB.seq:\*  
23: /cgn2\_6/ptodata/2/pna/US096A.COMB.seq:\*  
24: /cgn2\_6/ptodata/2/pna/US096B.COMB.seq:\*  
25: /cgn2\_6/ptodata/2/pna/US096C.COMB.seq:\*  
26: /cgn2\_6/ptodata/2/pna/US096D.COMB.seq:\*  
27: /cgn2\_6/ptodata/2/pna/US096E.COMB.seq:\*  
28: /cgn2\_6/ptodata/2/pna/US097A.COMB.seq:\*  
29: /cgn2\_6/ptodata/2/pna/US097B.COMB.seq:\*  
30: /cgn2\_6/ptodata/2/pna/US097C.COMB.seq:\*  
31: /cgn2\_6/ptodata/2/pna/US098A.COMB.seq:\*  
32: /cgn2\_6/ptodata/2/pna/US098B.COMB.seq:\*  
33: /cgn2\_6/ptodata/2/pna/US098C.COMB.seq:\*  
34: /cgn2\_6/ptodata/2/pna/US099A.COMB.seq:\*  
35: /cgn2\_6/ptodata/2/pna/US099B.COMB.seq:\*  
36: /cgn2\_6/ptodata/2/pna/US099C.COMB.seq:\*  
37: /cgn2\_6/ptodata/2/pna/US101.COMB.seq:\*  
38: /cgn2\_6/ptodata/2/pna/US101.COMB.seq:\*  
39: /cgn2\_6/ptodata/2/pna/US6000.COMB.seq:\*  
40: /cgn2\_6/ptodata/2/pna/US6001.COMB.seq:\*  
41: /cgn2\_6/ptodata/2/pna/US6002.COMB.seq:\*  
42: /cgn2\_6/ptodata/2/pna/US6003.COMB.seq:\*  
43: /cgn2\_6/ptodata/2/pna/US6004.COMB.seq:\*

44: /cgn2\_6/ptodata/2/pna/US6005.COMB.seq:\*  
45: /cgn2\_6/ptodata/2/pna/US6006.COMB.seq:\*  
46: /cgn2\_6/ptodata/2/pna/US6007.COMB.seq:\*  
47: /cgn2\_6/ptodata/2/pna/US6008.COMB.seq:\*  
48: /cgn2\_6/ptodata/2/pna/US6009.COMB.seq:\*  
49: /cgn2\_6/ptodata/2/pna/US6010.COMB.seq:\*  
50: /cgn2\_6/ptodata/2/pna/US6011.COMB.seq:\*  
51: /cgn2\_6/ptodata/2/pna/US6012.COMB.seq:\*  
52: /cgn2\_6/ptodata/2/pna/US6013.COMB.seq:\*  
53: /cgn2\_6/ptodata/2/pna/US6014.COMB.seq:\*  
54: /cgn2\_6/ptodata/2/pna/US6015.COMB.seq:\*  
55: /cgn2\_6/ptodata/2/pna/US6016.COMB.seq:\*  
56: /cgn2\_6/ptodata/2/pna/US6017.COMB.seq:\*  
57: /cgn2\_6/ptodata/2/pna/US6018.COMB.seq:\*  
58: /cgn2\_6/ptodata/2/pna/US6019.COMB.seq:\*  
59: /cgn2\_6/ptodata/2/pna/US6020.COMB.seq:\*  
60: /cgn2\_6/ptodata/2/pna/US6021.COMB.seq:\*  
61: /cgn2\_6/ptodata/2/pna/US6022.COMB.seq:\*  
62: /cgn2\_6/ptodata/2/pna/US6023.COMB.seq:\*  
63: /cgn2\_6/ptodata/2/pna/US6024.COMB.seq:\*  
64: /cgn2\_6/ptodata/2/pna/US6025.COMB.seq:\*  
65: /cgn2\_6/ptodata/2/pna/US6026.COMB.seq:\*  
66: /cgn2\_6/ptodata/2/pna/US6027.COMB.seq:\*  
67: /cgn2\_6/ptodata/2/pna/US6028.COMB.seq:\*  
68: /cgn2\_6/ptodata/2/pna/US6029.COMB.seq:\*  
69: /cgn2\_6/ptodata/2/pna/US6030.COMB.seq:\*  
70: /cgn2\_6/ptodata/2/pna/US6031.COMB.seq:\*  
71: /cgn2\_6/ptodata/2/pna/US6032.COMB.seq:\*  
72: /cgn2\_6/ptodata/2/pna/US6033.COMB.seq:\*  
73: /cgn2\_6/ptodata/2/pna/US6034.COMB.seq:\*  
74: /cgn2\_6/ptodata/2/pna/US6035.COMB.seq:\*  
75: /cgn2\_6/ptodata/2/pna/US6036.COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	20	US-09-536-784-1
2	1999	100.0	1999	30	US-09-765-271-1
3	1999	100.0	1999	30	US-09-765-272-1
4	1999	100.0	10383	41	US-60-029-960-81
5	1999	100.0	10711	13	US-08-961-527-145
6	1987.8	99.4	2160	22	US-09-583-110-1312
7	1987.8	99.4	2166	15	US-09-107-433-1102
8	1981.4	99.1	2160	1	PCT-US02-03987-9325
9	1981.4	99.1	2160	31	US-09-815-242-9325
10	1981.4	99.1	2160	37	US-10-072-851-9325
11	1981.4	99.1	10333	45	US-60-061-998-596
12	1631.2	81.6	9845	45	US-60-068-175-596
13	944.2	47.2	960	5	US-08-116-541-3
14	808.6	40.5	2172	1	PCT-US97-14436-139
15	808.6	40.5	2172	13	US-08-911-503-139
16	808.6	40.5	2172	13	US-08-911-503A-139
17	574.6	28.7	2337	1	PCT-US02-03987-6631
18	574.6	28.7	2337	37	US-09-815-242-6631
19	574.6	28.7	2337	37	US-10-072-851-6631
20	573	28.7	2370	15	US-09-134-000-1534
21	559	28.0	8395	14	US-09-070-927-217
22	559	28.0	8395	14	US-09-070-927A-217
23	553	27.7	2472	15	US-09-107-532-2013
24	553	27.7	2472	15	US-09-107-532A-2013
25	496.2	24.8	1386	37	US-10-091-007-91
26	410	20.5	2470	45	US-60-068-186-495
27	410	20.5	2474	44	US-60-050-444-495
28	352.2	17.6	3478	45	US-60-068-217-710
29	348.6	17.4	3279	24	US-09-634-238-44
30	348.6	17.4	7818	24	US-09-634-238-14
31	273	13.7	933	43	US-60-045-649-108



QY	1201	AATGACATACCAAGATTTTACAGTACACCAATGCAATTAACATAACACACCGAATCGA	1320
Db	1201	AATGACATACCAAGATTTTACAGTACACCAATGCAATTAACATAACACACCGAATCGA	1260
QY	1261	CAAAAAATATGAGACAGTAGTGAAGAAAGATGGCTGCTTACGCTGGCTTTGGCAAAATGG	1320
Db	1261	CAAAAAATATGAGACAGTAGTGAAGAAAGATGGCTGCTTACGCTGGCTTTGGCAAAATGG	1320
QY	1321	TGGAACTTACTATTAACCAATGTATATCCATAAAGTCGCTTTAGTGTAGTGGAGTGAAGA	1380
Db	1321	TGGAACTTACTATTAACCAATGTATATCCATAAAGTCGCTTTAGTGTAGTGGAGTGAAGA	1380
QY	1381	AGAGTTCCTCAATATGTGGAGACTGTGCCATGTGGAAGAAACGACAGCTATATGATGACCGA	1440
Db	1381	AGAGTTCCTCAATATGTGGAGACTGTGCCATGTGGAAGAAACGACAGCTATATGATGACCGA	1440
QY	1441	CATGATGAAAAACAGTCTTGTGACTATGGAACATGGACAAATGGCTATCTTGGTGGCTGCC	1500
Db	1441	CATGATGAAAAACAGTCTTGTGACTATGGAACATGGACAAATGGCTATCTTGGTGGCTGCC	1500
QY	1501	TCAGAGCTGTGTAATAACAGAACCTCTAATATACAGACGAGAAATTGAAAAACACATCAA	1560
Db	1501	TCAGAGCTGTGTAATAACAGAACCTCTAATATACAGACGAGAAATTGAAAAACACATCAA	1560
QY	1561	GACCTCTCAATTTGTGTACACCTATGATACATATTTGCTGGCTATACGCGTAATATTTCAAT	1620
Db	1561	GACCTCTCAATTTGTGTACACCTATGATACATATTTGCTGGCTATACGCGTAATATTTCAAT	1620
QY	1621	GGCTGTATGAGACAGGCTATTTCTAACCGCTCTGACACCACTGTGAGCAATAGGCTTACGGT	1680
Db	1621	GGCTGTATGAGACAGGCTATTTCTAACCGCTCTGACACCACTGTGAGCAATAGGCTTACGGT	1680
QY	1681	CGCTGCCAAAGTTTACCGCTCTATGTATGACCTACCTGTCTGGAAGACCAATCCAGAAGA	1740
Db	1681	CGCTGCCAAAGTTTACCGCTCTATGTATGACCTACCTGTCTGGAAGACCAATCCAGAAGA	1740
QY	1741	TTGGAATATATACAGAGGGGCTCTACAGAAATGAGAGATTCGTATTTAAAAATGCTGCTCG	1800
Db	1741	TTGGAATATATACAGAGGGGCTCTACAGAAATGAGAGATTCGTATTTAAAAATGCTGCTCG	1800
QY	1801	TTTACGCTGGGAACACACCTGCTCCACAAACCCCATCACTGAAAGTTCAGAGCTCATC	1860
Db	1801	TTTACGCTGGGAACACACCTGCTCCACAAACCCCATCACTGAAAGTTCAGAGCTCATC	1860
QY	1861	ATCAGATAGTTCAACTTACAGACGTCTAGCTCAACACACTCCACACACAAATATATAGTAGAC	1920
Db	1861	ATCAGATAGTTCAACTTACAGACGTCTAGCTCAACACACTCCACACACAAATATATAGTAGAC	1920
QY	1921	TACCAATCTCAACAAATATATACGCAACAAATCAATACACCCCTGATCAACAAATACAGAA	1980
Db	1921	TACCAATCTCAACAAATATATACGCAACAAATCAATACACCCCTGATCAACAAATACAGAA	1980
QY	1981	TTCTCAACACGACACACCA 1999	
Db	1981	TTCTCAACACGACACACCA 1999	
RESULT 2			
US-09-765-271-1			
; Sequence 1, Application US/09765271			
; GENERAL INFORMATION:			
; APPLICANT: Choi et. al.			
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines			
; NUMBER OF SEQUENCES: 452			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Human Genome Sciences, Inc.			
; STREET: 9410 Key West Avenue			
; CITY: Rockville			
; STATE: Maryland			
; COUNTRY: USA			
; ZIP: 20850			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage			

```

1      COMPUTER:  HP Vectra 486/33
2      OPERATING SYSTEM:  MSDOS version 6.2
3      SOFTWARE:  ASCII text
4
5      CURRENT APPLICATION DATA:
6      APPLICATION NUMBER:  US/09/765,271
7      FILING DATE:  22-Jan-2001
8      CLASSIFICATION:  <unknown>
9
10     PRIOR APPLICATION DATA:
11     APPLICATION NUMBER:  09/536,784
12     FILING DATE:  <unknown>
13     APPLICATION NUMBER:  08/961,083
14     FILING DATE:  OCT-30-1997
15
16     ATTORNEY/AGENT INFORMATION:
17     NAME:  Michelle S. Marks
18     REGISTRATION NUMBER:  41,971
19     REFERENCE/DOCKET NUMBER:  PB340P3
20
21     TELECOMMUNICATION INFORMATION:
22     TELEPHONE:  (301) 309-8504
23     TELEFAX:  (301) 309-8512
24
25     INFORMATION FOR SEQ ID NO: 1:
26
27     SEQUENCE CHARACTERISTICS:
28
29     LENGTH: 1999 base pairs
30     TYPE: nucleic acid
31     STRANDEDNESS: double
32     TOPOLOGY: linear
33
34     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
35
36     US-09-765-271-1

```

Query Match	Similarity	100.0%	Score 1999:	DB 30;	Length 1999:
Matches	Conservative	0;	Mismatches	0;	Indels
					Gaps
Qy	1	TAAATCTACGCAATATAAAATCAACTATTTGCTGACTTGGTCTTGAAAGCGCGCTCAA	60		
Db	1	TAAATCTACGCAATATAAAATCAACTATTTGCTGACTTGGTCTTGAAAGCGCGCTCAA	60		
Qy	61	TGCCCAAGCTATGATTTTCCCAAGATTTGGTTAAGCAATCGTTTCTATGGAAGACCA	120		
Db	61	TGCCCAAGCTATGATTTTCCCAAGATTTGGTTAAGCAATCGTTTCTATGGAAGACCA	120		
Qy	121	TGCGTCTTCGACACACAGGGGATTTGATACATCCGTATCTGGGAGCTTTCTTGCGCAA	180		
Db	121	TGCGTCTTCGACACACAGGGGATTTGATACATCCGTATCTGGGAGCTTTCTTGCGCAA	180		
Qy	181	TCTGCAAGCAATTCCTCTCCACAGGTGATCAACTCTCACCAACAGTTGATTAAGTTGAC	240		
Db	181	TCTGCAAGCAATTCCTCTCCACAGGTGATCAACTCTCACCAACAGTTGATTAAGTTGAC	240		
Qy	241	TTACTTTTCACTTGACACTTCGACACAGCTATTTCTGTATAGGCTCAGGAAGCTTGGTT	300		
Db	241	TTACTTTTCACTTGACACTTCGACACAGCTATTTCTGTATAGGCTCAGGAAGCTTGGTT	300		
Qy	301	AGCGATTCAGTTAGAGCAAAAAGACACAGCAGACAGAAATCTTGACTACTATATAATTA	360		
Db	301	AGCGATTCAGTTAGAGCAAAAAGACACAGCAGAAATCTTGACTACTATATAATTA	360		
Qy	361	GGTCTACATGTCCTAATGGGAAGTATGGAATGAGACAGCAGCTCATAAATCTACTATGTA	420		
Db	361	GGTCTACATGTCCTAATGGGAAGTATGGAATGAGACAGCAGCTCATAAATCTACTATGTA	420		
Qy	421	AGACCTCATATTTAAGTTTACGTACAGTTAGCTGCTGGGTGGAATCCCTCAGGACAC	480		
Db	421	AGACCTCATATTTAAGTTTACGTACAGTTAGCTGCTGGGTGGAATCCCTCAGGACAC	480		
Qy	481	AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCCAAGACGCGCGAAACTTGGCTT	540		
Db	481	AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCCAAGACGCGCGAAACTTGGCTT	540		
Qy	541	ATTCGAATATGAAAAATCAAGGCTAATCTCTCTGTAACAGTTTGAGAAAGCATATATAC	600		
Db	541	ATTCGAATATGAAAAATCAAGGCTAATCTCTCTGTAACAGTTTGAGAAAGCATATATAC	600		



D	b	1	TAAATCTACGACAACTAAANAATCAACTCACTGCTGACTTGGGTTCTGAACGGCCGGCTCA	60
O	y	61	TGCCCAAGCTATGATATATTCGCCACAGATTTGGTTAAGCAATCGTTTCTATGGAAGACA	120
D	b	61	TGCCCAAGCTATGATATATTCGCCACAGATTTGGTTAAGCAATCGTTTCTATGGAAGACA	120
O	y	121	TGCGTTCTTGACACACAGGGGGGATTTGATACATCCGTATCCTGGGAGCTTCTTGCCGA	180
D	b	121	TGCGTTCTTGACACACAGGGGGGATTTGATACATCCGTATCCTGGGAGCTTCTTGCCGA	180
O	y	181	TCTGCAAGCAATATTCCTCCAGAGTGGATCACTCCACCAACAGTTGATTAACTTAC	240
D	b	181	TCTGCAAGCAATATTCCTCCAGAGTGGATCACTCCACCAACAGTTGATTAACTTAC	240
O	y	241	TTACTTTTCAACTTCGACTTCGACCAAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT	300
D	b	241	TTACTTTTCAACTTCGACTTCGACCAAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT	300
O	y	301	AGCGATTCACTTTAGACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360
D	b	301	AGCGATTCACTTTAGACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360
O	y	361	GGCTACATGCTCTAATGGGAAGTATGGAAATGCAAGAGAGTCAAAACTCTATATGTA	420
D	b	361	GGCTACATGCTCTAATGGGAAGTATGGAAATGCAAGAGAGTCAAAACTCTATATGTA	420
O	y	421	AGACCTCAATATTTAAGTTTACCTCAAGTTAGCTTGCTGCTGGAAATGCTCAGGACAC	480
D	b	421	AGACCTCAATATTTAAGTTTACCTCAAGTTAGCTTGCTGCTGGAAATGCTCAGGACAC	480
O	y	481	AAACCATATATGACCCCTATTACATCCAGAGCAGCCCAAGCCGCCAAGCTTGCTT	540
D	b	481	AAACCATATATGACCCCTATTACATCCAGAGCAGCCCAAGCCGCCAAGCTTGCTT	540
O	y	541	ATTCGAATGAAAAATCAAGGCTCAATCTGCTGAAGCTATGAAAGCAAGCTCAATAC	600
D	b	541	ATTCGAATGAAAAATCAAGGCTCAATCTGCTGAAGCTATGAAAGCAAGCTCAATAC	600
O	y	601	ACCAATTAAGTATGACTACAAAGTCTCAAAATGAGCAACTAATTAACCTGCTTACATGGA	660
D	b	601	ACCAATTAAGTATGACTACAAAGTCTCAAAATGAGCAACTAATTAACCTGCTTACATGGA	660
O	y	661	TAAATTAACCTCAAGGAGTCAATCAATCAAGTTGAGAAGAAACAGGCTATTAACCTACTAC	720
D	b	661	TAAATTAACCTCAAGGAGTCAATCAATCAAGTTGAGAAGAAACAGGCTATTAACCTACTAC	720
O	y	721	AACGGGATGGATGTCACAAANTGTAGACCAAGAGTCAAAAACATCTGGGATAT	780
D	b	721	AACGGGATGGATGTCACAAANTGTAGACCAAGAGTCAAAAACATCTGGGATAT	780
O	y	781	TTAATAATACAGACAATACCTTGCTCATCCAGATGCAATTTGCAAGTGCCTTCAACAT	840
D	b	781	TTAATAATACAGACAATACCTTGCTCATCCATGCCATGCAATTTGCAAGTGCCTTCAACAT	840
O	y	841	TGTTGATGTTTTCTAAGGTTAAAGTCATTTGCCAGCTAGGAGCAGCCATCACTCAAGTAA	900
D	b	841	TGTTGATGTTTTCTAAGGTTAAAGTCATTTGCCAGCTAGGAGCAGCCATCACTCAAGTAA	900
O	y	901	TGTTTCCCTGGGAATTAACCAAGCAGATGAAGAAACAAACGCGACTGGGATCAACTATGAA	960
D	b	901	TGTTTCCCTGGGAATTAACCAAGCAGATGAAGAAACAAACGCGACTGGGATCAACTATGAA	960
O	y	961	ACCGATCACAGACTATGCTCTGCTGCTGGAGTACGCTGATCACTCACTGCTACTAT	1020
D	b	961	ACCGATCACAGACTATGCTCTGCTGCTGGAGTACGCTGATCACTCACTGCTACTAT	1020
O	y	1021	CGTTACAGATGAGCCCTATATCTACCTCGGAGCAAAATACCTCTGTTTAACTGGGATAG	1080
D	b	1021	CGTTACAGATGAGCCCTATATCTACCTCGGAGCAAAATACCTCTGTTTAACTGGGATAG	1080
O	y	1081	GGGCTACTTTGGCAACATCACTTGCAANTCGCCCTGCAACATGGCGAAAGTCCGACG	1140
D	b	1081	GGGCTACTTTGGCAACATCACTTGCAANTCGCCCTGCAACATGGCGAAAGTCCGACG	1140

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

```

1  RESULT      4
2  US-60-029-960-81/c
3  : Sequence 81, Application US/60029960
4  : GENERAL INFORMATION:
5  : APPLICANT: Charles Kunsch
6  : TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
7  : NUMBER OF SEQUENCES: 1643
8  : CORRESPONDENCE ADDRESS:
9  : ADDRESSEE: Human Genome Sciences, Inc.
10 : STREET: 9410 Key West Avenue
11 : CITY: Rockville
12 :

```

STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/60/029,960  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340DP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 81:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10383 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-60-029-960-81

Query Match 100.0%; Score 1999; DB 41; Length 10383;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TAAATCTACGCAATTAATAATCAATCTATTGCTGACTTGGTTCGAGCCCGCTCAA 60  
DB 9763 TAAATCTACGCAATTAATAATCAATCTATTGCTGACTTGGTTCGAGCCCGCTCAA 9704  
OY 61 TCCCAAGCTAATGATATTCCTCCAGATTTGGTTAAGCAATCGTTTCATGGAAGACCA 120  
DB 9703 TCCCAAGCTAATGATATTCCTCCAGATTTGGTTAAGCAATCGTTTCATGGAAGACCA 9644  
OY 121 TCGCTTCTCGACACAGGGGGATTGATACATCCGTATCTGGAAGCTTCTTCCGCAA 180  
DB 9643 TCGCTTCTCGACACAGGGGGATTGATACATCCGTATCTGGAAGCTTCTTCCGCAA 9584  
OY 181 TCTGCAAGCAATTCCTTCGCAAGGTGATCACTCTACCCAAAGTTGATTAACTGAC 240  
DB 9583 TCTGCAAGCAATTCCTTCGCAAGGTGATCACTCTACCCAAAGTTGATTAACTGAC 9524  
OY 241 TTACTTTCAATGAGACTTCGACAGACTATTTCTGTAGGGCTCAGGAAGCTTGGT 300  
DB 9523 TTACTTTCAATGAGACTTCGACAGACTATTTCTGTAGGGCTCAGGAAGCTTGGT 9464  
OY 301 AGCGATTGAGTAGAACAACCAAGCAAGCAAGAAATCTGACTACTATATAATA 360  
DB 9463 AGCGATTGAGTAGAACAACCAAGCAAGCAAGAAATCTGACTACTATATAATA 9404  
OY 361 GGTCTACATGCTTAATGGAACTATGAATGAGACAGAGCTCAAAAATCTATGTGA 420  
DB 9403 GGTCTACATGCTTAATGGAACTATGAATGAGACAGAGCTCAAAAATCTATGTGA 9344  
OY 421 AGACTCAATATTAAGCTTACCTCAGTTAGCTTGTGGCTGGAATGCCCAAGCAC 480  
DB 9343 AGACTCAATATTAAGCTTACCTCAGTTAGCTTGTGGCTGGAATGCCCAAGCAC 9284  
OY 481 AAACCAATATGACCCCTATTTCATCATCAGAAAGACCCCAAGAACTTGGTCT 540  
DB 9283 AAACCAATATGACCCCTATTTCATCATCAGAAAGACCCCAAGAACTTGGTCT 9224  
OY 541 ATCTGAAATGAAATCAAGGCTACATCTGTGCTGAACAGTATGAAAGCAGTCAATAC 600  
|||||

DB 9223 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAAAGCAGTCAATAC 9164  
OY 601 ACCAATTAAGTAGGAGCTACCAAGTCTCAATACAGCAAGTAAATACCTGCTTACATGGA 660  
DB 9163 ACCAATTAAGTAGGAGCTACCAAGTCTCAATACAGCAAGTAAATACCTGCTTACATGGA 9104  
OY 661 TAATTAAGTCAAGGAAGTCAATCAAGTGAAGGAAGGAAGCAAGGCTTAATACCTACTAC 720  
DB 9103 TAATTAAGTCAAGGAAGTCAATCAAGTGAAGGAAGGAAGCAAGGCTTAATACCTACTAC 9044  
OY 721 AACTGGATGAGTGTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGGGATAT 780  
DB 9043 AACTGGATGAGTGTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGGGATAT 8984  
OY 781 TTACATACAGAGCAATAGCTTGCCTATCCAGACAGATGAATTCGAAGTCTTACCAT 840  
DB 8983 TTACATACAGAGCAATAGCTTGCCTATCCAGACAGATGAATTCGAAGTCTTACCAT 8924  
OY 841 TGTGATGTTTCTAAAGGTAAGTCAATGCTCCAGCTAGAGACAGCCATCACTCAATGA 900  
DB 8923 TGTGATGTTTCTAAAGGTAAGTCAATGCTCCAGCTAGAGACAGCCATCACTCAATGA 8864  
OY 901 TGTTCCTTCCGAATTAACCAAGCAGTAGAAACCAACCCGACCTGGGGATCAACTATGAA 960  
DB 8863 TGTTCCTTCCGAATTAACCAAGCAGTAGAAACCAACCCGACCTGGGGATCAACTATGAA 8804  
OY 961 ACCGATACAGACTATGCTCCTGCTGGAGTACGCTGTCTAGATTCAGTCACTACTAT 1020  
DB 8803 ACCGATACAGACTATGCTCCTGCTGGAGTACGCTGTCTAGATTCAGTCACTACTAT 8744  
OY 1021 CGTTCAGATGAGCCCTATTAATACCTGCGGACAAATACCTCTGTTAATACCTGGGATAG 1080  
DB 8743 CGTTCAGATGAGCCCTATTAATACCTGCGGACAAATACCTCTGTTAATACCTGGGATAG 8684  
OY 1081 GGGCTACTTGGCAACATCTACCTTGCATACGCTCGCAACAATGGCCAAAGCTCCAC 1140  
DB 8683 GGGCTACTTGGCAACATCTACCTTGCATACGCTCGCAACAATGGCCAAAGCTCCAC 8624  
OY 1141 CGTGAACCTTAACAAGGTGGAGCTCAACCGCCCAAGACTTTCCTTAATAGTCTAG 1200  
DB 8623 CGTGAACCTTAACAAGGTGGAGCTCAACCGCCCAAGACTTTCCTTAATAGTCTAG 8564  
OY 1201 AATGACTACCCAGATTAATCACTCAATGCCATTTCAATTAACACACCGAATCAGA 1260  
DB 8563 AATGACTACCCAGATTAATCACTCAATGCCATTTCAATTAACACACCGAATCAGA 8504  
OY 1261 CAAAAAATATGAGCAAGTAGAAGTAAGTGGCTGCTGCTTACGCTTGGCAAAATGG 1320  
DB 8503 CAAAAAATATGAGCAAGTAGAAGTAAGTGGCTGCTGCTTACGCTTGGCAAAATGG 8444  
OY 1321 TGAAGCTTACTTAACCAATGATATATGCAATGAAGTGTCTGCTTACGCTTGGCAAAATGG 1380  
DB 8443 TGAAGCTTACTTAACCAATGATATATGCAATGAAGTGTCTGCTTACGCTTGGCAAAATGG 8384  
OY 1381 AGAGTCTCTAATGTGGAAGTCTGCAATGGAAGAAACGACAGCTATATGATGACGA 1440  
DB 8383 AGAGTCTCTAATGTGGAAGTCTGCAATGGAAGAAACGACAGCTATATGATGACGA 8324  
OY 1441 CATGATGAAGAGTCTTACTATGGAAGTGAAGAAATGCTATCTTGGTGGCTCC 1500  
DB 8323 CATGATGAAGAGTCTTACTATGGAAGTGAAGAAATGCTATCTTGGTGGCTCC 8264  
OY 1501 TCAGGCTGGTAACAGGAACTTCACTATCAACAGGGAATTTGAAACCAATCA 1560  
DB 8263 TCAGGCTGGTAACAGGAACTTCACTATCAACAGGGAATTTGAAACCAATCA 8204  
OY 1561 GACCTCTCAATTTTGAACCTGATGAATTTCTGCTGCTTAATGCGCTAAATATTCAT 1620  
DB 8203 GACCTCTCAATTTTGAACCTGATGAATTTCTGCTGCTTAATGCGCTAAATATTCAT 8144  
OY 1621 GGTGTATGAGAGGCTATTTCAACGCTGTGACACCACTTGTAGGCAATGGCTTACGCT 1680  
DB 8143 GGTGTATGAGAGGCTATTTCAACGCTGTGACACCACTTGTAGGCAATGGCTTACGCT 8084  
|||||



QY 1681 CGCTGCCAAGTTTACCGCTATGAGCTTACCTGTCTGGAAGGAAGCAATCCAGAGA 1740  
|||||  
Db 8083 CGCTGCCAAGTTTACCGCTATGAGCTTACCTGTCTGGAAGGAAGCAATCCAGAGA 8024  
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAATTTAAAAATGGTCTCG 1800  
|||||  
Db 8023 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAATTTAAAAATGGTCTCG 7964  
QY 1801 TTCTAGTGGAACTCAGCTGCTCCACAAACACCCCATCACTGAAAGTTCAAGCTCAGC 1860  
|||||  
Db 7963 TTCTAGTGGAACTCAGCTGCTCCACAAACACCCCATCACTGAAAGTTCAAGCTCAGC 7904  
QY 1861 ATCAGATATGTAACCTTCAAGTCTGATGCTCAACCACTCCAGCAGCAATATATATGTCGAC 1920  
|||||  
Db 7903 ATCAGATATGTAACCTTCAAGTCTGATGCTCAACCACTCCAGCAGCAATATATGTCGAC 7844  
QY 1921 TACCAATCTTAACAATATATAGCAACAATCAATATACACCCCTGATCAACAATAATGAGAA 1980  
|||||  
Db 7843 TACCAATCTTAACAATATATAGCAACAATCAATATACACCCCTGATCAACAATAATGAGAA 7784  
QY 1981 TTCTCAACGAGCAACCA 1999  
|||||  
Db 7783 TTCTCAACGAGCAACCA 7765

RESULT 5  
US-08-961-527-145/C  
; Sequence 145, Application US/08961527  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: P3340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 145:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10711 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-961-527-145

Query Match 100.0%; Score 1999; DB 13; Length 10711;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATAAAAATCAACTCATGTGCTGACTTGGGTCTGGAAGCGCGGTCAA 60

Db 9767 TAAATCTACGACAAATAAAAATCAACTCATGTGCTGACTTGGGTCTGGAAGCGCGGTCAA 9708  
|||||  
QY 61 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGCAATCGTTTCATGCAAGACCA 120  
|||||  
Db 9707 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGCAATCGTTTCATGCAAGACCA 9648  
QY 121 TGCTTCTTCGACACAGGGGGATTTATACATCCGTATTCCTGGAGCTTCTTGGCGCA 180  
|||||  
Db 9647 TGCTTCTTCGACACAGGGGGATTTATACATCCGTATTCCTGGAGCTTCTTGGCGCA 9588  
QY 181 TCGCAAGCAATTTCCCTCAAGAGGTGATATCACTTCACCCAAACATGTAATTAAGTTGAC 240  
|||||  
Db 9587 TCGCAAGCAATTTCCCTCAAGAGGTGATATCACTTCACCCAAACATGTAATTAAGTTGAC 9528  
QY 241 TTACTTTTCAACTTGCAGCTTCGACGACGACTATTTTCGTAAAGCTTCAGGAAGCTTGGTT 300  
|||||  
Db 9527 TTACTTTTCAACTTGCAGCTTCGACGACGACTATTTTCGTAAAGCTTCAGGAAGCTTGGTT 9468  
QY 301 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTGACCTACTATATTAATAA 360  
|||||  
Db 9467 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTGACCTACTATATTAATAA 9408  
QY 361 GGTCTACATGCTATATGGAACATATGGAATGCAGACAGCTCAAAACTACTATGTTAA 420  
|||||  
Db 9407 GGTCTACATGCTATATGGAACATATGGAATGCAGACAGCTCAAAACTACTATGTTAA 9348  
QY 421 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCTTGTGGCTGGAATCCCTCAGGCAAC 480  
|||||  
Db 9347 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCTTGTGGCTGGAATCCCTCAGGCAAC 9288  
QY 481 AAACCAATATGACCCCTATTTACATCCAGACAGCCCAAGACCGCGCAAACTTGGTCTT 540  
|||||  
Db 9287 AAACCAATATGACCCCTATTTACATCCAGACAGCCCAAGACCGCGCAAACTTGGTCTT 9228  
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTGCTGTAACGTTGAGAAAGCAAGTCAATAC 600  
|||||  
Db 9167 ATCTGAAATGAAAAATCAAGGCTACATCTGCTGTAACGTTGAGAAAGCAAGTCAATAC 9108  
QY 601 ACCAATTTAGTATGATGACATCAAAAGTCTCAATTCAGCAAGTAATTTACCTGCTTACATGA 660  
|||||  
Db 9167 ACCAATTTAGTATGATGACATCAAAAGTCTCAATTCAGCAAGTAATTTACCTGCTTACATGA 9108  
QY 661 TAATTAACCTCAAGGAAGTATATCAATTAAGTTGAAGAAAGCAAGGCTATTAACCTACTGAC 720  
|||||  
Db 9107 TAATTAACCTCAAGGAAGTATATCAATTAAGTTGAAGAAAGCAAGGCTATTAACCTACTGAC 9048  
QY 721 AACTGGGATGATGCTTACACAAATGTAGACCAAGAAAGCTCAAAAAATCTGTGGATAT 780  
|||||  
Db 9047 AACTGGGATGATGCTTACACAAATGTAGACCAAGAAAGCTCAAAAAATCTGTGGATAT 8988  
QY 781 TTACAATACAGACGAATATACGTTTGCCTATTCAGACGATGAATTCGAAGTGGCTTACCAT 840  
|||||  
Db 8987 TTACAATACAGACGAATATACGTTTGCCTATTCAGACGATGAATTCGAAGTGGCTTACCAT 8928  
QY 841 TGTGTATGTTTCTTAACGTAAGTATATGCTCCAGCTAGAGACAGCCGATCAATCAAGTAA 900  
|||||  
Db 8927 TGTGTATGTTTCTTAACGTAAGTATATGCTCCAGCTAGAGACAGCCGATCAATCAAGTAA 8868  
QY 901 TGTTCCTTCGGAATTTAACCAAGCAGTAGAACAACCCGAGCTGGGGATCAACTATGAA 960  
|||||  
Db 8867 TGTTCCTTCGGAATTTAACCAAGCAGTAGAACAACCCGAGCTGGGGATCAACTATGAA 8808  
QY 961 ACCGATACAGACTATGCTGCTGCTTGGAGTACGGTGTCTACGATCAACTGCTACTAT 1020  
|||||  
Db 8807 ACCGATACAGACTATGCTGCTGCTTGGAGTACGGTGTCTACGATCAACTGCTACTAT 8748  
QY 1021 CGTTACGATGAGCCCTATTAACCTACCCCTGGACAAATATCTCTGTTTATTAATGGGATAG 1080  
|||||  
Db 8747 CGTTACGATGAGCCCTATTAACCTACCCCTGGACAAATATCTCTGTTTATTAATGGGATAG 8688  
QY 1081 GGGTACTTTGGCAACATACCTTGCAATACGCTGCAACAAATGCGCAAGAGTCCCAAC 1140  
|||||

Db	8687	GGGCTACTTGGCAACATCATCCCTTGCATTAACGGCTTGCAACCAATCGGAACGTCGCCAGC	86528
Qy	1141	CGTGGAAACTCTAAACAAAGTGGGACTCAACCGCCGCAAGACTTTCTAAATGTGCTAGG	1200
Db	8627	CGTGGAAACTCTAAACAAAGTGGGACTCAACCGCCGCAAGACTTTCTAAATGTGCTAGG	8568
Qy	1201	AATCGACTCCCAAGATTCTACTACTCAATGGCATTTTCAAGTAAACAACACCGAATTCAGA	1260
Db	8567	AATCGACTCCCAAGATTCTACTACTCAATGGCATTTTCAAGTAAACAACACCGAATTCAGA	8508
Qy	1261	CAAAAAATATGAGACAAAGTAGTGAAGAAGTGGCTGGCTTACGCTGCTTTGCAATATG	1320
Db	8507	CAAAAAATATGAGACAAAGTAGTGAAGAAGTGGCTGGCTTACGCTGCTTTGCAATATG	8448
Qy	1321	TGGAACTTCTATTAACAATGATATATCCATTAAGTGTCTTTAGTGAAGAGTGAATA	1380
Db	8447	TGGAACTTCTATTAACAATGATATATCCATTAAGTGTCTTTAGTGAAGAGTGAATA	8388
Qy	1381	AAGATTCTTATGTGCGAAGTGTGCGCATGTGAAGGAAGACAGCGCTATATGTATGCCGA	1440
Db	8387	AAGATTCTTATGTGCGAAGTGTGCGCATGTGAAGGAAGACAGCGCTATATGTATGCCGA	8328
Qy	1441	CATGATGAAACAGTCTTACTTATGTGAAGTGGACCAAAATCCCTATCTGCTGGCTCCC	1500
Db	8327	CATGATGAAACAGTCTTACTTATGTGAAGTGGACCAAAATCCCTATCTGCTGGCTCCC	8268
Qy	1501	TCAGGCTGTGTAACAGAGAACTCTTACTATATACAGACGAGGAATTTGAAAAACCATCA	1560
Db	8267	TCAGGCTGTGTAACAGAGAACTCTTACTATATACAGACGAGGAATTTGAAAAACCATCA	8208
Qy	1561	GACCTCTCAATTTGTAGCACCTGATGAACATATTTGCTGGCTATAGCGGTAAATATTCAT	1620
Db	8207	GACCTCTCAATTTGTAGCACCTGATGAACATATTTGCTGGCTATAGCGGTAAATATTCAT	8148
Qy	1621	GGCTGATGAGAGGCTATTTCAACCGTCAACACCACTGTGTGGGAATGGGCTTACGGT	1680
Db	8147	GGCTGATGAGAGGCTATTTCAACCGTCAACACCACTGTGTGGGAATGGGCTTACGGT	8088
Qy	1681	CGCTGCCAAAGTTTACCGCTCTATGATGACCTACTGCTGTGAAGGAAGCAATCCAGAGA	1740
Db	8087	CGCTGCCAAAGTTTACCGCTCTATGATGACCTACTGCTGTGAAGGAAGCAATCCAGAGA	8028
Qy	1741	TTTGAATATACCAAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGGTGGTG	1800
Db	8027	TTTGAATATACCAAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGGTGGTG	7968
Qy	1801	TTTCTGCTGAGAACTCAACCGTCCACACAACCCCATCACTCAAGTGAAGTTCAAGGTATC	1860
Db	7967	TTTCTGCTGAGAACTCAACCGTCCACACAACCCCATCACTCAAGTGAAGTTCAAGGTATC	7908
Qy	1861	ATCAGATAGTTCAACTTCAAGTCTAGCTCAACCACTCCAAAGCAGACAATATATGTAGCAG	1920
Db	7907	ATCAGATAGTTCAACTTCAAGTCTAGCTCAACCACTCCAAAGCAGACAATATATGTAGCAG	7848
Qy	1921	TACCAATCTTAACATATATAGCAACACATCAAAATACACCCCTGATCAACAAATATAGAA	1980
Db	7847	TACCAATCTTAACATATATAGCAACACATCAAAATACACCCCTGATCAACAAATATAGAA	7788
Qy	1981	TCCTCAACGACGACAACCA 1999	
Db	7787	TCCTCAACGACGACAACCA 7769	

```

RESULT      6
US-09-583-110-1312
: Sequence 1312, Application US/09583110
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucelte-Stamm et al.
: TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
: TITLE OF INVENTION: Pneumoniae for Diagnosis and Therapeutics
: FILE REFERENCE: PPTH00-07A
: CURRENT APPLICATION NUMBER: US/09/583, 110
: CURRENT FILING DATE: 2000-05-26

```

```

; PRIOR APPLICATION NUMBER: US 09/1107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 1312
; TYPE: DNA
; LENGTH: 2160
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-1312

Query Match      99.4%; Score 1987.8; DB 22; Length 2160;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY   1 TAAATCTACGCAATAAATCATCACTCTTCCTACTTGGGTGCAGAGCGCGTGCA 60
    |||||||
Db   159 taaatctacgcaataaaaatcaactctctgtcacttgggttcctgaagcgcgctcaa 218

OY   61 TCCCAAGCTATGATATTCCACCAGATTGGTTAAGGCAATCGTTTTATCGAAGACA 120
    |||||||
Db   219 tgcccaagctaattgatattcccacgatttggttaaggaaacgltttcatlagaagaca 278

OY   121 TCGCTTCTTCGACACACAGGGGGATTGATACCATCCGATCTCTGGAGACTTCTTGCCCAA 180
    |||||||
Db   279 tcgcttcttcgcaccaaggggagttagatcacaccgtatccctcgtggagcttcttcgcgaa 338

OY   181 TCTGGAAGAAGCATTCCTCCTCCAGGTGGATGCATCTCACGCCAACAGTTGATTAACTTGC 240
    |||||||
Db   339 ctctgaagaagcaatccctccccaaglygatacactcactcaacagtlgattaaagtgac 398

OY   241 TTACTTTTCAACTTGTGACACTTCGACACAGACTATTCTCGTAAGCGTCAGAGACTTGTT 300
    |||||||
Db   399 ttacttttcaacttgcacttcgccacgaatacttctcgttaaagctcaggaagcttggt 458

OY   301 AGCGATTGACGTTAGAACAAAAAGAACCAAGCAAGAAATCTTGACCTACTATTAATTA 360
    |||||||
Db   459 agcgatttcgattagacaanaaagaacaaccaagaagaatacttgacctactataataaa 518

OY   361 GGCTACATGTTTAATGGGAATATMGAAATGACAGACAGAGCTCAAAACTACTATAGTGA 420
    |||||||
Db   519 ggctacatglttaatlgygaactatgyaatgaagacgaagctcaaaactactatcylgtaa 578

OY   421 ACACCTGCATATATTTAAGTTTACTCTCAGTAGTACCTTTGGCTGGGAATGCTCAGACACC 480
    |||||||
Db   579 agacctcaataattaaagtttacccagttagccttgctgctgtaaygctcccaagcacc 638

OY   481 AAACCAATATGACCCCTATTTCACATCCAGAACAGCCCCAAGACCGCCGAACCTTGCTTT 540
    |||||||
Db   639 aaaccaatatgccccctatttcacalcgaagcagcgcccaagccgcggaacttggctt 698

OY   541 ATCTGAATGAAAATTCAAAGGCTAATCTCTGCTGAACAGTATGAGAAAGCATGAATAC 600
    |||||||
Db   699 atctgaatgaaaaatcaaggtctaacctctcgcylgaacaagltgynaagcylgtaaatc 758

OY   601 ACCAATTTCTGATGACATCAAGATCTCAAAATCAGCAAGTAATTCACCGTCTACATGGA 660
    |||||||
Db   759 accaatctctgctggtctacaagaagctcctaatacgaagtaattatccctgcttaaciyga 818

OY   661 TAATTACCTCAAGGAAGTCATCATCAAGTTGAGAGAAAGAAACAGGCTTAACCTACTGAC 720
    |||||||
Db   819 taattacctcaaggaagtcataatcaatcaagtlgaagaagaacagcctataacctactcac 878

OY   721 AACTGGGATGGATGTTCTACACAATGTATGACCAAGACAGCTCAAAAACATCTGTGGGATAT 780
    |||||||
Db   879 aactggatggtatgtctacacaaatgtagaccaaagaagctcbaaaaactctgtgysalat 938

OY   781 TTACAAATACAGAGCAATTCGTTGCTATTCAGAGATGAATTCGAAGTCGCTTCACAT 840
    |||||||
Db   939 ttacaatcacagcgaatcgttgcctacccaagaagatgaattggaagtcgcttccacat 998

```







```
DB 2139 tccttaaccagcacacca 2157
|||||
RESULT 9
US-09-815-242-9325
: Sequence 9325, Application US/09815242
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21*
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9325
: LENGTH: 2160
: TYPE: DNA
: ORGANISM: Streptococcus pneumoniae
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(2160)
US-09-815-242-9325

Query Match 99.18; Score 1981.4; DB 31; Length 2160;
Best Local Similarity 99.48; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATTCAGACAAATAAATCAATGACTGTGGTCTGTAAGCGCGGTCAA 60
DB 159 taatatccagcaataaataacatcactgactggtctgagcgcgcgtcaa 218
QY 61 TGCCCAAGCTATGATATTCACAGATTTGTTAAGCAATGTTTCTATCGAAGCA 120
DB 219 tgcacaagcataigatatcccaacagatttggttaagcaatcglttctacagaagca 278
QY 121 TCGCTTCTTGACCAAGAGGGGATTTGATACATCCGTAATCTTGAGAGCTTTTGGCGAA 180
DB 279 tcgcttctcgaacacaggggagtgtgataccatccgtatcctcggaggtcttctcgcaaa 338
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCAACTCAACCCCAAGTGAATTAGTTGAC 240
DB 339 tcgtcaagaacatccctccaagglgagtcctcctcaacaaagttgatalaagttgac 398
QY 241 TTATTTTCACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCGAAGGCTTGTT 300
DB 399 ttaactttcaacttcgacttcgacacagactatcttcgttaaggtcctcgaaggttggtc 458
QY 301 AGCATTTCACTTGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATTAATAA 360
|||||
```

```
DB 459 agcgatcagctagacaaaaagcaaccaagcaagaatcttgacctataataataa 518
QY 361 GGTCTACATGCTCTAATGGGAACTATGGAATGCACACAGACCTCAAAATCTATTGGTAA 420
DB 519 ggtcacaatgcttaatgggaactatggaaatgcagacagactcaaaactactatgtaa 578
QY 421 AGACCTCAATATTTAAGTTTACCTCAGTTAGCCCTTGCTGGCTGGAATGCTTCAGGCACC 480
DB 579 agacctcaataatttaagtttaacctagtttagccttgcttgcttggaatgctcgaagcc 638
QY 481 AAACCAATATGACCCCTATTTCATCCAGAAAGACCCCAACACCGCGAATCTTGCTT 540
DB 639 aaaccaatataagccacctataccaacgaagcagcccaagaccgcgcgaacttgctct 698
QY 541 ATCTGAATGAATAATCAAGGCTCATCTGCTGAACAGATGAGAAAGCAAGTCAATATAC 600
DB 699 atcggaaatgaataaatacaaggctacacccgcgcgaaagatgagaagaagccaataac 758
QY 601 ACCAATTAAGTATGAGCTACAAAGTCTCAAAATCAGCAAGTAAATTAACCTGCTTACATGA 660
DB 759 accaatctatgattggtctacaagaatctcaaatcgaagtaattacaccttgcttacaatga 818
QY 661 TAATTACTCAAGAAAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTTAATCACTATCAC 720
DB 819 taattacctcaaggaagtcataccaagttaagaaagaacaggtctataacctactcac 878
QY 721 AACTGGGATGGAATGCTCAACAAATGTAAGACCAAGGCTCAAAACACTGCTGGGATAT 780
DB 879 aactggagatgagtcctcaacaacaatgtagaccaagaagctcaaaaacatcgtggtgatat 938
QY 781 TTACAATACACAGCAATACGTTGCTTATCCAGACAGATGAATTCAGTGGCTTGTACCAT 840
DB 939 ttacaataacagagaataacgttgctctacccaagatgaattgtaagtcgtcttacaat 998
QY 841 TGTGATGTTTCTAAGGTAAGTCAATTTGCCACCTAGAGACCACTATAGTAAAGTAA 900
DB 999 tgtgattgtttctaaagctaaagtcattgcccagctagagcaagcatcagtaagtaa 1058
QY 901 TGTTCCTGCGAATTAACCAAGCAGTAGAACAACCGGACGTGGGATCAATATGAA 960
DB 1059 tgttcctcctggaaatcaacaagaagtagaacaacaacgcgacgtgggatacctaataatgaa 1118
QY 961 ACCGATACACAGACTATGCTCTGCTTGGAGTACGGTGTCTACGATTCACATGCTACTAT 1020
DB 1119 acgatacacagacatgctccttgcttgagtagcgtgctcagatgcaactgacactat 1178
QY 1021 CGTTCAGATGATGACCTTATACCTGAGGACAAATATCTCTTTATTAATGGGATAG 1080
DB 1179 cgttcacgattgagccctatactacccctgggacaataacccctgtttataacttggatag 1238
QY 1081 GGGCTACTTGGCAACATCACTTGCATATACGCCCTGCAACATCGGGAACGTCCAGC 1140
DB 1239 gggctactttggcaacatcaatccttgcaataacgcccctgcaacaatcggaaaagctcccaagc 1298
QY 1141 CGTGAATACTCTAAACAAGTTCGAGCTCAACCGCGCCCAAGCTTTCTTAATGGTCTAGC 1200
DB 1299 cgtgaaacatctaaacaagtcgagacccaacgcgccaagcttctctaattggtctcgg 1358
QY 1201 AATGACTACCCAGTATTCATCTACTCAATATGCCATTTCAAGTAAACCAACGAATAGA 1260
DB 1359 aatgactaccacaagattactactactcaaatgccaattccaagtaacaacacgaataga 1418
QY 1261 CAATAAATATGAGACAGTGTGAAAGATGAGGCGTGTAGCGCTTGCTTGGCAATGG 1320
DB 1419 caaataaataatgagcaagtcagtgaaaagaatggtgcgtcttaacgctccttgcaaatgg 1478
QY 1321 TGAATCTTACTATTAACCAATGTATATCAATAAAGTGCTTTTGTGATGAGGAGTGAATA 1380
DB 1479 tgaacttactataataacaatgatatccataaagtcgtctcttagtgatggagtgaaaa 1538
QY 1381 AGAGTTCTTATATGTCGAACCTGTCGCATGAGAGGAACGACAGCTTATATGATGACCGA 1440
DB 1539 agagttctctaattcgtgaactcgtgcataatgaagaagcaagcagcctatatgatgaccca 1598
|||||
```











```

ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PM-0009-2 P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-416
INFORMATION FOR SEO ID NO: 596:
SEQUENCE CHARACTERISTICS:
LENGTH: 9845 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: genomic DNA
IMMEDIATE SOURCE:
CLONE: SPN2c598
US-60-068-175-596

Query Match      81.6%; Score 1631.2; DB 45; Length 9845;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 1776; Conservative 0; Mismatches 223; Indels 1; Gaps 1;

QY 1 TAAATCTACGACAAATAAATACTCATCTGACTTGGTTCTGAACGCCGCTCAA 60
DB 9734 TAAATCTACGACAAATAAATACTCATCTGACTTGGTTCTGAACGCCGCTCAA 9675

QY 61 TGCCCAAGCTAATATATATCCACAGATTTGGTTAAGCAATCTTTCTATCAGAACCA 120
DB 9674 TGCCCAAGCTAATATATATCCACAGATTTGGTTAAGCAATCTTTCTATCAGAACCA 9615

QY 121 TCGCTTTCGACCAAGAGGGGATTTGATACCATCCGATCCGAGAGCTTCTTGCGCAA 180
DB 9614 TCGCTTTCGACCAAGAGGGGATTTGATACCATCCGATCCGAGAGCTTCTTGCGCAA 9555

QY 181 TCTGCAAGCAATTCCTCCAGAGTGATGATCACTCCACACAGTTGATTAAAGTTAC 240
DB 9554 TCTGCAAGCAATTCCTCCAGAGTGATGATCACTCCACACAGTTGATTAAAGTTAC 9495

QY 241 TTACTTTTCAACTTCGACTTCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300
DB 9494 TTACTTTTCAACTTCGACTTCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 9435

QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTATATATAATTA 360
DB 9434 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTATATATAATTA 9375

QY 361 GGTTCATATGCTTAATGGGAATATGAAATGACAGACAGCTTAATAACTATATGTTAA 420
DB 9374 GGTTCATATGCTTAATGGGAATATGAAATGACAGACAGCTTAATAACTATATGTTAA 9315

QY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGGTGGCTGAAGTGGCTCAGGGACC 480
DB 9314 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGGTGGCTGAAGTGGCTCAGGGACC 9235

QY 481 AAACCAATATGACCCCTATTTCATCCAGAGAGCCCAAGCCGCGAAACTTGGTCTT 540
DB 9234 AAACCAATATGACCCCTATTTCATCCAGAGAGCCCAAGCCGCGAAACTTGGTCTT 9135

QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATAGAAAAGCATATAC 600
DB 9194 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATAGAAAAGCATATAC 9135

QY 601 ACCAATTAAGTGGACTACAAAGTCTCAATTCAGCAAGTAAATTAACCT- GCTTACATGG 659
DB 9134 ACCAATTAAGTGGACTACAAAGTCTCAATTCAGCAAGTAAATTAACCTGCTTACATGG 9075

QY 660 ATAATTAAGTCAAGAGTATCATCAATCAAGTTGAAGAAAGCAAGCTATTAACCTACTCA 719
DB 9074 ATAATTAAGTCAAGAGTATCATCAATCAAGTTGAAGAAAGCAAGCTATTAACCTACTCA 9015

QY 720 CAATGGAGATGATGCTACCAAAATGTAGCCAAAGAGCTCAAAAACATCTGTGGGATA 779
DB 720 CAATGGAGATGATGCTACCAAAATGTAGCCAAAGAGCTCAAAAACATCTGTGGGATA 779

```

```

DB 9014 CTACTGGAGTATGTTTACAAATGTAGACCAAGAAAGCTCAAAAACATCTGTGGGATA 8955
QY 780 TTACAAATACAGACAAATACGTTGCTATCCAGACGATGAATTCGAAGTCCCTTACCA 839
DB 8954 TCTACAACTCCGATCAATACGTTGCTATCCAGACGATGAATTCGAAGTCCCTTACCA 8895
QY 840 TTGTTGATGTTTCAACGATTAAGTCAATTCGCCAGTATGAGACAGCCCATCATGATGTA 899
DB 8894 TCGTAGATGTTTCAAAATGATGAATGATCAGCCCACTGGAGCTGTCACCAACAGTA 8835
QY 900 ATGTTTCTTGGGAATTAACCAACGATGAGAAACAAACCCGGAGTGGGATCACTATGA 959
DB 8834 ACGTTTATTTTGTACCAACCAACGATGAGAAACCAACGATGAGTGGGATTCCTATGA 8775
QY 960 AACGATACAGACTATGCTCCTGAGTACGAGTGTCTACATTCACCTACTACTA 1019
DB 8774 AACCAATCAGCGATTAATGACACTGCTGATAGATAGGAGTTATGATTTCCATCAGACTA 8715
QY 1020 TCGTTACAGATGAGCCCTATATACCTGAGACAAATCTCTGTTTAAATACGAGATA 1079
DB 8714 TCGTTAATGATATTCCTTATTAATCCGGAGACAGACACTGTCTACAACTGGGATA 8655
QY 1080 GGGCTACTTTGGCAATCACCCTTGCATATCGCCCTGCACAAATCCGGAAGCTCCAG 1139
DB 8654 GAGCATATTTGCTATATATATCTGCAATATGCTCTTCAACCAATCAGAAATGTCACAG 8595
QY 1140 CCGTGGAACTCTAACAAGTGGGACTCAACCGCGCAACACTTCTTAATAGTGTAG 1199
DB 8594 CCGTTGAGACTTGAATTAAGTGGGCTGATGATGATGATGATGATGATGATGATGATG 8535
QY 1200 GAATGACTACCAAGTATCTACTACTACTACTACTACTACTACTACTACTACTACTACT 1259
DB 8534 GTATGACTATTCACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8475
QY 1260 ACAAATAATATGAGCAAGTATGTAAGAAATGAGTGTCTGCTTACGCTGCTTCAATG 1319
DB 8474 AATAACAATATGAGCAAGTATGTAAGAAATGAGTGTCTGCTTACGCTGCTTCAATG 8415
QY 1320 GTGAGACTTACTATTAACCAATGATATATCCATTAAGTGTGTTAGTATGATGAGTAAA 1379
DB 8414 GTGAGACTTACTATTAACCAATGATATATCCATTAAGTGTGTTAGTATGATGAGTAAA 8355
QY 1380 AAGATTTCTCTAATGTGGAAGTGTGCTGATGAGAAAGCAAGCTATATATGATGACCG 1439
DB 8354 AAGATTTCTCTAATGTGGAAGTGTGCTGATGAGAAAGCAAGCTATATATGATGACCG 8295
QY 1440 ACATGATGAAGAACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1499
DB 8294 ACATGATGAAGAACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 8235
QY 1500 CTCAGGCTGATGAAGAAAGCAAGCTCTAATCTATACAGAGAGAAATGAAAACCATCA 1559
DB 8234 CTCAGGCTGATGAAGAAAGCAAGCTCTAATCTATACAGATGAGAAAGTGAAGAAACCATCA 8175
QY 1560 AGACCTCTCAATTTGTAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
DB 8174 AGACCTCTCAATTTGTAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 8115
QY 1620 TGGCTGATGAGCAAGGCTATTTCTAACGCTGACACACACTGTATGAGCAATGGCTTACG 1679
DB 8114 TGGCTGATGAGCAAGGCTATTTCTAACGCTGATGATGATGATGATGATGATGATGATGAT 8055
QY 1680 TCGCTGCAAAAGTGTACCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739
DB 8054 TCGCTGCAAAAGTGTACCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7995
QY 1740 ATTGGAATATACAGAGAGGCTCTACAGAAATGAGAAATTCGATTTAAATAGTGTCTC 1799
DB 7994 ACTGAGAGATGCGCAGAGGACTTTTCAGAAAGCGGGAATTTGATTCAGAAATAGGACTC 7935
QY 1800 GTTCTACGATGAGTACCTGCTGACACACACCCCATCATCACTGAAAGTTCAAGCTCAT 1859
DB 7934 GCCCAATATGAGTGAACCCCTTACTCAACATCTCAACAGCTGGAAGTTCAAGCTCAT 7875

```



APPLICANT: Stodola, Robert  
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,  
POLYPEPTIDES AND THEIR USES  
NUMBER OF SEQUENCES: 713  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/14436  
FILING DATE: 15-AUG-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/024,022  
FILING DATE: 16-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50533  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 139:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2172 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
PCT-US97-14436-139

Query Match 40.5%; Score 808.6; DB 1; Length 2172;  
Best Local Similarly 99.4%; Pred. No. 5.1e-218;  
Matches 822; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1174 CGCGAAGCTTCTCTAATGCTAGGATGCTACCCAGTATCTCACTCAATAATGC 1233  
1 CGCGAAGCTTCTCTAATGCTAGGATGCTACCCAGTATCTCACTCAATAATGC 60  
DB 1234 CATTTCAGTAACACACGCAATCAGACAAAAAATATGAGCAAGTAGTGAAGATGCG 1293  
61 CATTTCAGTAACACACGCAATCAGACAAAAAATATGAGCAAGTAGTGAAGATGCG 120  
QY 1294 TGCCTGCTTACGCTCCCTTTGCAAAATGCTGCACTTACTATTAACCAATGATATCATAT 1353  
121 TGCCTGCTTACGCTCCCTTTGCAAAATGCTGCACTTACTATTAACCAATGATATCATAT 180  
DB 1354 AGTGTCTTTAGTATGATGAGTGAAGAAAGTCTCTAATGTGGAAGTGTGCGCATGAA 1413  
181 AGTGTCTTTAGTATGATGAGTGAAGAAAGTCTCTAATGTGGAAGTGTGCGCATGAA 240  
QY 1414 GGAAGACAGACGCTATATGATGACGACATGATGAAAAACAGTCTTGAATATGCAACTGCG 1473  
241 GGAAGACAGACGCTATATGATGACGACATGATGAAAAACAGTCTTGAATATGCAACTGCG 300  
DB 1474 ACG-AAATGCTATCTGCTTGGCTCCCTCAGCTGTGTAAGACGAACTCTAATATA 1552  
301 ACGAAATGCTATCTGCTTGGCTCCCTCAGCTGTGTAAGACGAACTCTAATATA 360  
QY 1553 CAGACGAGAAATGTAAGACACATCAAGACCTCTCAATTTGTGACACCTGATCAACTAT 1592  
361 CAGACGAGAAATGTAAGACACATCAAGACCTCTCAATTTGTGACACCTGATCAACTAT 420

QY 1593 TTGCTGCTATATACGCGTAATATTCATATGCTGTATGAGACAGGCTATTTCAACCGTCTGA 1652  
421 TTGCTGCTATATACGCGTAATATTCATATGCTGTATGAGACAGGCTATTTCAACCGTCTGA 480  
QY 1653 CACCCTTTGTAAGCAATGGCTTACGGTCCGCAAGTTTACCCTCTATGTATGACT 1712  
481 CACCCTTTGTAAGCAATGGCTTACGGTCCGCAAGTTTACCCTCTATGTATGACT 540  
QY 1713 ACCTGTGGAAGGAAAGCAATCAGAGATTTGGAATATACGAGGGGCTCTACAGAAATG 1772  
541 ACCTGTGGAAGGAAAGCAATCAGAGATTTGGAATATACGAGGGGCTCTACAGAAATG 600  
QY 1773 GAGAAATCGTATTTAAATAGTCTGCTTCTAGCTGGAACCTCACTCTCCACAAAC 1832  
601 GAGAAATCGTATTTAAATAGTCTGCTTCTAGCTGGAACCTCACTCTCCACAAAC 660  
QY 1833 CCCCATCACTGAAAGTTTCAAGCTCATCATAGTATGTTCACTTACAGCTTACGCTCAA 1892  
661 CCCCATCACTGAAAGTTTCAAGCTCATCATAGTATGTTCACTTACAGCTTACGCTCAA 720  
QY 1893 CCACCCAGCACAATAATATAGTACGACTACCAATCCTAACAAATATACGACAAATCAA 1952  
721 CCACCCAGCACAATAATATAGTACGACTACCAATCCTAACAAATATACGACAAATCAA 780  
QY 1953 ATACACCCCTGATCAACAAATCAGAAATCCTCAACGACGACAAACCA 1999  
781 ATACACCCCTGATCAACAAATCAGAAATCCTCAACGACGACAAACCA 827  
DB

RESULT 15  
US-08-911-503-139  
Sequence 139, Application US/08911503  
GENERAL INFORMATION:  
APPLICANT: Black, Michael  
APPLICANT: Hodgson, John  
APPLICANT: Knowles, David  
APPLICANT: Lonetto, Michael  
APPLICANT: Nicholas, Richard  
APPLICANT: Stodola, Robert  
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,  
POLYPEPTIDES AND THEIR USES  
NUMBER OF SEQUENCES: 713  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: IBM Compatible  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/911,503  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/024,022  
FILING DATE: 16-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50533  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 139:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2172 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-08-911-503-139

Query Match 40.5%; Score 808.6; DB 13; Length 2172;  
Best Local Similarity 99.4%; Pred. No. 1e-218;  
Matches 822; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1174 CGCCAAGACTTCTCTAAATGCTAGGAATCGACTACCCAGTATCTACTACTCAAAATGC 1233  
DB 1 CGCCAAGACTTCTCTAAATGCTAGGAATCGACTACCCAGTATCTACTACTCAAAATGC 60  
QY 1234 CATTTCAGTAAACACACACCGAATTCAGACAAAAATATGAGCAGTATGGAAGATGGC 1293  
DB 61 CATTTCAAGTAACACACCGAATTCAGACAAAAATATGAGCAGTATGGAAGATGGC 120  
QY 1294 TGGCTGCTTACGCTGCTTGGCAATGCTGAACTTACTATAACCATGTATATCCATTA 1353  
DB 121 TGGCTGCTTACGCTGCTTGGCAATGCTGAACTTACTATAACCATGTATATCCATTA 180  
QY 1354 AGTCGCTTTAGTAGTGGAGTGAAGAAAGTCTCTAATGTGGAACCTGTCATGAA 1413  
DB 181 AGTCGCTTTAGTAGTGGAGTGAAGAAAGTCTCTAATGTGGAACCTGTCATGAA 240  
QY 1414 GGAACGACAGCCTATATGATGACCGACATGATGAAAACAGTCTTGACTTATGAACTGG 1473  
DB 241 GGAACGACAGCCTATATGATGACCGACATGATGAAAACAGTCTTGACTTATGAACTGG 300  
QY 1474 ACG-AAATGCTATCTTGTGTGGCTCCCTCAGGCTGGTAAACAGGAACCTTAATA 1532  
DB 301 ACGAAATGCTATCTTGTGTGGCTCCCTCAGGCTGGTAAACAGGAACCTTAATA 360  
QY 1533 CAGACGAGGAATTTGAAAACCATCAAGACCTCTCAATTTGTAGACCTGATGAACCTAT 1592  
DB 361 CAGACGAGGAATTTGAAAACCATCAAGACCTCTCAATTTGTAGACCTGATGAACCTAT 420  
QY 1593 TTGCTGGCTATACGCGTAAATATTCATGCGTATGACAGGCTATTTCTAACCGTCTGA 1652  
DB 421 TTGCTGGCTATACGCGTAAATATTCATGCGTATGACAGGCTATTTCTAACCGTCTGA 480  
QY 1653 CACCACTTGTAGCAATGGCCTTAGCGTGGCTGCCAAAGTTTACCGCTCTATGATGACCT 1712  
DB 481 CACCACTTGTAGCAATGGCCTTAGCGTGGCTGCCAAAGTTTACCGCTCTATGATGACCT 540  
QY 1713 ACCTGCTGAAGAGCAATCCAGAAGATTGGAATATACAGAGGGGCTCTACAGAAATG 1772  
DB 541 ACCTGCTGAAGAGCAATCCAGAAGATTGGAATATACAGAGGGGCTCTACAGAAATG 600  
QY 1773 GAGAAATTCGTAATTTAAAAATGGTGTCTGCTTACGTGAACCTGCTCCACAACAAC 1832  
DB 601 GAGAAATTCGTAATTTAAAAATGGTGTCTGCTTACGTGAACCTGCTCCACAACAAC 660  
QY 1833 CCCCATCAACTGAAGTTAAGCTCATCTCATGATAGTTCAACTTACAGTCTAGCTCAA 1892  
DB 661 CCCCATCAACTGAAGTTAAGCTCATCTCATGATAGTTCAACTTACAGTCTAGCTCAA 720  
QY 1893 CCACCTCAAGCAACAATATATAGACACTACCAATCCCTAACATATATAGCAACATCAA 1952  
DB 721 CCACCTCAAGCAACAATATATAGACACTACCAATCCCTAACATATATAGCAACATCAA 780  
QY 1953 ATACAACCCCTGATCAACAAATCAGAAATCTCAACCCAGACACCA 1999  
DB 781 ATACAACCCCTGATCAACAAATCAGAAATCTCAACCCAGACACCA 827

Search completed: August 11, 2002, 04:16:37  
Job time: 9723 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 11, 2002, 02:09:49 ; Search time 442.87 seconds  
(without alignments)  
11011.969 Million cell updates/sec

Title: US-08-961-083-1  
1999  
Sequence: 1 TAAATCTACGACATATAAA.....ATCGTCACACGACACCA 1999

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1437365 seqs, 1219827628 residues  
Total number of hits satisfying chosen parameters: 2874730

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents\_NA\_New:\*  
1: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:\*  
7: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq2:\*  
8: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	4	US-08-961-083-1
2	1989	100.0	10711	7	US-10-158-844-145
3	117.2	5.9	9100	7	US-10-121-120-27
4	59.2	3.0	2370	5	US-09-540-209B-3719
5	57.8	2.9	7571	7	US-10-158-844-2
6	55.2	2.8	428	5	US-09-673-476-206
7	51.8	2.6	363	5	US-09-673-476-626
8	44.4	2.2	1104	7	US-10-179-131-3144
9	44.2	2.2	1350	7	US-10-179-131-2857
10	40.6	2.0	223	5	US-09-673-476-438
11	40.4	2.0	1511	7	US-10-179-131-3015
12	40	2.0	610	7	US-10-027-632-5967
13	39.2	2.0	867	5	US-09-216-393B-340
14	39.2	2.0	1397	5	US-09-216-393B-342
15	39.2	2.0	1397	5	US-09-216-393B-343
16	39.2	2.0	1397	5	US-09-216-393B-345
17	38.6	1.9	1143	7	US-10-179-131-1548
18	38	1.9	430	7	US-10-027-632-40693
19	37.6	1.9	2805	7	US-10-155-881-23712
20	37.4	1.9	242	5	US-09-789-189-1228
21	37.4	1.9	2034	7	US-10-179-131-3855
22	37.2	1.9	1611	8	US-60-360-039-46474
23	36.6	1.8	679	7	US-10-027-632-213202
24	36.6	1.8	2409	8	US-60-360-039-46149
25	36.6	1.8	8093	7	US-10-172-086-26

c 26	36.4	1.8	473	5	US-09-673-476-671	Sequence 671, App
c 27	36.4	1.8	9021	7	US-10-172-086-54	Sequence 54, Appl
c 28	36	1.8	689	7	US-10-027-632-148124	Sequence 148124,
c 29	35.8	1.8	348	5	US-09-673-476-99	Sequence 99, Appl
c 30	35.6	1.8	445	5	US-09-918-995-15709	Sequence 15709, A
c 31	35.6	1.8	2373	5	US-09-540-209B-4136	Sequence 4136, Ap
c 32	35.4	1.8	8045	5	US-09-913-878A-1	Sequence 1, Appl1
c 33	35.2	1.8	1812	7	US-10-179-131-2414	Sequence 2414, Ap
c 34	35.2	1.8	3234	8	US-60-385-568-185	Sequence 185, App
c 35	35.2	1.8	3402	7	US-10-179-131-1603	Sequence 1603, Ap
c 36	35	1.8	510	7	US-10-027-632-82144	Sequence 82144, A
c 37	35	1.8	132205	5	US-09-578-519B-25	Sequence 309095,
c 38	35	1.8	132205	5	US-09-578-243B-25	Sequence 25, Appl
c 39	35	1.8	132205	5	US-09-578-518B-25	Sequence 25, Appl
c 40	35	1.8	132205	5	US-09-579-262B-25	Sequence 25, Appl
c 41	35	1.8	132205	5	US-09-579-262B-25	Sequence 25, Appl
c 42	35	1.8	132205	7	US-10-185-047-25	Sequence 25, Appl
c 43	35	1.8	132205	7	US-10-185-047A-25	Sequence 25, Appl
c 44	34.8	1.7	1655	7	US-10-175-525-86	Sequence 86, Appl
c 45	34.8	1.7	2709	7	US-10-179-131-3751	Sequence 3751, Ap

## ALIGNMENTS

```

RESULT 1
US-08-961-083-1
; Sequence 1, Application US/08961083
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OR INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8512
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-961-083-1

Query Match 100.0%; Score 1999; DB 4; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Gaps 0;
QY 1 TAAATCTACGACATATAAAATCAGCTGCTGACTTGCGTTTCGACGCGGCTCA 60

```

|||||  
Db 1 TAAATCTAGACAAATAAATCAATCTGAGCTTGGGTTCTGAAGCGCCGCCAA 60  
QY 61 TGGCAAGTAATGATTTCCACAGATTTGGTTAAAGCAATGTTTCTATTCGAAGACA 120  
Db 61 TGGCCAAAGCTAATGATTTCCACAGATTTGGTTAAAGCAATGTTTCTATTCGAAGACA 120  
QY 121 TGGCTTCTTGACACAGGGGGATTTGATACATCCGTATCCCTGGAGCTTCTTGCGCAA 180  
Db 121 TGGCTTCTTGACACAGGGGGATTTGATACATCCGTATCCCTGGAGCTTCTTGCGCAA 180  
QY 181 TCGCAAGAAATTTCCCTCCAGGTGATGATCTTCACCCACAGTGTATTAAGTTGAC 240  
Db 181 TCGCAAGAAATTTCCCTCCAGGTGATGATCTTCACCCACAGTGTATTAAGTTGAC 240  
QY 241 TTAATTTTCACTTTCGACTTCGACAGACTATTTCTCGTAAGGCTCAGAGAGCTTGGT 300  
Db 241 TTAATTTTCACTTTCGACTTCGACAGACTATTTCTCGTAAGGCTCAGAGAGCTTGGT 300  
QY 301 AGCGATTCAGTTAGAACAAAGCAACAGCAAGAAATCTTGACCTATATATAATA 360  
Db 301 AGCGATTCAGTTAGAACAAAGCAACAGCAAGAAATCTTGACCTATATATAATA 360  
QY 361 GGTCTACATCTCTAATGGGAATCTATGGAATGCAGACAGCTCAAAACTATATGTTAA 420  
Db 361 GGTCTACATCTCTAATGGGAATCTATGGAATGCAGACAGCTCAAAACTATATGTTAA 420  
QY 421 AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTGGCTGGAGTGGCTCAGGACC 480  
Db 421 AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTGGCTGGAGTGGCTCAGGACC 480  
QY 481 AAACCAATATGACCCCTATTACATCCAGAGAGCCCAAGCCGCGAAACTTGGCTT 540  
Db 481 AAACCAATATGACCCCTATTACATCCAGAGAGCCCAAGCCGCGAAACTTGGCTT 540  
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACATATGAGAAACACTCAATAC 600  
Db 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACATATGAGAAACACTCAATAC 600  
QY 601 ACCAATTAAGTATGATGATCAAGAGTCAATAGCAAGTAATTAACCTCTTACATAGA 660  
Db 601 ACCAATTAAGTATGATGATCAAGAGTCAATAGCAAGTAATTAACCTCTTACATAGA 660  
QY 661 TAAATACCTCAGAGAAATCATCAATCAAGTTGAAGAAAGAAAGGCTATTAACCTACTAC 720  
Db 661 TAAATACCTCAGAGAAATCATCAATCAAGTTGAAGAAAGAAAGGCTATTAACCTACTAC 720  
QY 721 AACTGGGATGATGATCTACCAAAATGTAAGCAAGAAAGTCAAAACATCTGCGGATAT 780  
Db 721 AACTGGGATGATGATCTACCAAAATGTAAGCAAGAAAGTCAAAACATCTGCGGATAT 780  
QY 781 TTAACAATACAGAGAAATGATGCTATCCAGAGATGATGATGATGATGATGATGATGAT 840  
Db 781 TTAACAATACAGAGAAATGATGCTATCCAGAGATGATGATGATGATGATGATGATGAT 840  
QY 841 TGTGTATGTTTCTAAGCGTAAGTATGCGCAGCTAGAGACAGCCATCAGTCAAGTAA 900  
Db 841 TGTGTATGTTTCTAAGCGTAAGTATGCGCAGCTAGAGACAGCCATCAGTCAAGTAA 900  
QY 901 TGTTCCTCTCGGAATTAACCAAGCAATGAGAAACCGGAGCTGGGATCAACTATGAA 960  
Db 901 TGTTCCTCTCGGAATTAACCAAGCAATGAGAAACCGGAGCTGGGATCAACTATGAA 960  
QY 961 ACCGATCACAGACTATGCTGCTGCTGAGTGGAGTGTACTAGATTCAACCTGCTACTAT 1020  
Db 961 ACCGATCACAGACTATGCTGCTGCTGAGTGGAGTGTACTAGATTCAACCTGCTACTAT 1020  
QY 1021 CGTTACAGATGAGCCCTATTAACCTCGGACAAATACTCTGTTTATTAACCTGGGATAG 1080  
Db 1021 CGTTACAGATGAGCCCTATTAACCTCGGACAAATACTCTGTTTATTAACCTGGGATAG 1080  
QY 1081 GGGCTACTTTGGCAATCATCCTTGCAATACGCCCTGCAACAATCGGAAACGCTCCAGC 1140  
|||||

Db 1081 GGGCTACTTTGGCAATCATCCTTGCAATAGCCCTGCAACAATCGGAAACGCTCCAGC 1140  
QY 1141 CGTGAAGAACTTAACAAAGGTGAGACTCAACCCGCGCAAGACTTCTTAATAGTCTAGG 1200  
Db 1141 CGTGAAGAACTTAACAAAGGTGAGACTCAACCCGCGCAAGACTTCTTAATAGTCTAGG 1200  
QY 1201 AATGACTACCAAGTATTCAGTCAATCAATGCAATTTCAAGTAAACCAACCAAGTACGA 1260  
Db 1201 AATGACTACCAAGTATTCAGTCAATCAATGCAATTTCAAGTAAACCAACCAAGTACGA 1260  
QY 1261 CAAAAAATATGAGCAAGTATGAAAAAGAGTGGCTGCTAGGCTGCTTTCGCAATAGG 1320  
Db 1261 CAAAAAATATGAGCAAGTATGAAAAAGAGTGGCTGCTAGGCTGCTTTCGCAATAGG 1320  
QY 1321 TGGAACTTACTATTAACCAATGATATATCCATTAAGTGTCTTTAGTATGAGAGTGAAGA 1380  
Db 1321 TGGAACTTACTATTAACCAATGATATATCCATTAAGTGTCTTTAGTATGAGAGTGAAGA 1380  
QY 1381 AGAGTTCTCTAATGTGGAATCTGTGCATGANGAAGACAGCCTATATGATGACCGA 1440  
Db 1381 AGAGTTCTCTAATGTGGAATCTGTGCATGANGAAGACAGCCTATATGATGACCGA 1440  
QY 1441 CATGATGAAGAAAGCTCTGCTATGTAAGTGAACGCAAAATGCTATCTTGGCTGCC 1500  
Db 1441 CATGATGAAGAAAGCTCTGCTATGTAAGTGAACGCAAAATGCTATCTTGGCTGCC 1500  
QY 1501 TCAGGCTGCTAAGAAAGCAAGCACTCTAATCTATACAGAGAGAAATTTGAACCAACATCA 1560  
Db 1501 TCAGGCTGCTAAGAAAGCAAGCACTCTAATCTATACAGAGAGAAATTTGAACCAACATCA 1560  
QY 1561 GACCTCTCAATTTGTAGACACTGATGAATATTTGCTGGCTATACCGTAAATATTCAT 1620  
Db 1561 GACCTCTCAATTTGTAGACACTGATGAATATTTGCTGGCTATACCGTAAATATTCAT 1620  
QY 1621 GGTCTATGAGACAGGCTATCTAACCCTGACACCACTGTGAGGAATGGCTTACGGT 1680  
Db 1621 GGTCTATGAGACAGGCTATCTAACCCTGACACCACTGTGAGGAATGGCTTACGGT 1680  
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGAAAGCAATCCAGAA 1740  
Db 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGAAAGCAATCCAGAA 1740  
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTTGTAATTAATAGTGGCTCG 1800  
Db 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTTGTAATTAATAGTGGCTCG 1800  
QY 1801 TTCTAGTGAATCTACACCTGCTCCACAAACCCCATCAACCTGAAGTTCAAGCTCATC 1860  
Db 1801 TTCTAGTGAATCTACACCTGCTCCACAAACCCCATCAACCTGAAGTTCAAGCTCATC 1860  
QY 1861 ATCAGATAGTTCAACTTCAAGCTAGCTCAACCACTCCAGACCAAAATATAGTACGAC 1920  
Db 1861 ATCAGATAGTTCAACTTCAAGCTAGCTCAACCACTCCAGACCAAAATATAGTACGAC 1920  
QY 1921 TACCAATCTTAACCAATATATGCAACAAATCAATCAACCCCTGATCAACAAATCAGAA 1980  
Db 1921 TACCAATCTTAACCAATATATGCAACAAATCAATCAACCCCTGATCAACAAATCAGAA 1980  
QY 1981 TCCTCAACGACGACCAACA 1999  
Db 1981 TCCTCAACGACGACCAACA 1999

RESULT 2  
US-10-158-844-145/c  
; Sequence 145, Application US/10158844  
; GENERAL INFORMATION:  
; APPLICANT: Kunsch et al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue











```

; SEQ ID NO 2857
;
; LENGTH: 1350
;
; TYPE: DNA
;
; ORGANISM: Candida albicans
;
US-10-179-131-2857

```

Query Match	2.2%;	Score 44.2;	DB 7;	Length 1350;
Best Local Similarity	60.3%;	Pred. No. 0.052;	Mismatches 48;	Indels 0;
Matches 73;	Conservative	0;	Mismatches	48;

Qy	1846	AACTCAAGTCTGATTCATTCAGATAGTTCACATTCACAGTCTACTCTCAACACATCCACAGAC	1905
Db	986	aaagatcaacacacacacacacacaaataatcaaatccaatccaatccaatccaatccaagaagcc	1045
Qy	1906	AAATTAATAGTAGACATACCAATTCCTACAACTTAATACGACACATCAATACACCCCTGA	1965
Db	1046	aaataaataagcaccacacacacacacagttatccagttatccacacacacacacacacacagcagcaaa	1105
Qy	1966	T 1966	
Db	1106	t 1106	

```

: RESULT 10
: US-09-673-476-438
: Sequence 438, Application US/09673476
: GENERAL INFORMATION:
: APPLICANT: STEWART
: APPLICANT: BOCHRIESEN-BROSCH, ROLAND
: APPLICANT: COLE, STEPHEN
: APPLICANT: GORDON, STEPHEN
: APPLICANT: BILLAULT, ALAIN
: TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
: TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
: TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
: FILE REFERENCE: 05394.0011-00000
: CURRENT APPLICATION NUMBER: US/09/673,476
: CURRENT FILING DATE: 2002-03-29
: PRIOR APPLICATION NUMBER: PCT/IB99/00740
: PRIOR FILING DATE: 1999-04-16
: PRIOR APPLICATION NUMBER: 09/060,756
: PRIOR FILING DATE: 1998-04-16
: NUMBER OF SEQ ID NOS: 743
: SOFTWARE: PatentIn Ver. 2.2
: SEQ ID NO 438
: LENGTH: 223
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (136)
: OTHER INFORMATION: a, t, c or g
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (158)
: OTHER INFORMATION: a, t, c or g
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (203)
: OTHER INFORMATION: a, t, c or g
: US-09-673-476-438

```

Query Match	2.0%	Score 40.6	DB 5	Length 223
Best Local Similarity	48.6%	Pred. No. 0.28		
Matches 106	Conservative 0	Mismatches 112	Indels 0	Gaps 0

  

QY	281	AAGCTCAGGAACTGGTTAGCAATTAGTTAGAACAAAAAGCAACCAAGCAAGAATAATC	340
Db	6	aagcttcgagatccgagatgcacacgcctgcgacaaagacaccttcacaaatctcyaaatc	65
	341	TTGACCTACTATTATMAATAGAGTCTACATGTCCTAATGGAACTATGGAATGCAGACACA	400

Db 66 ctgacccgatactcgtgaacccgctgctgttcggcacaataccctcgttcgagcgatgacgagcgcg 125

QY 401 GCTCAAAACTACTATGGTTAAAGACCTCAATATATTTAAAGTTTACCTGATTAGCCTTGCTG 460

Db 126 ggcgaacagcttcttcgcatcaacgcgtccgancgtgatctgacgaacgcgagcgctcgtg 185

QY 461 GCTGAATGCTCTGGCCACCAACCAATATGACCCCTTA 498

Db 186 gcgcgcatgctgacatcmaccacgacgctcaaccgcta 223

```

RESULT 11
US-10-179-131-3015
Sequence 3015, Application US/10179131
GENERAL INFORMATION:
APPLICANT: HARE, ROBERTA S.
APPLICANT: SHAW, KAREN J.
APPLICANT: SHIMER JR., GEORGE H.
APPLICANT: KESSLER, MARCO
APPLICANT: NOLLING, JORK
APPLICANT: ZENG, QIANDONG
APPLICANT: GREENE, JONATHAN R.
TITLE OF INVENTION: CANDIDA ALBICANS NUCLEIC ACIDS AND POLYPEPTIDES,
FILE REFERENCE: 2976-4031
CURRENT FILING DATE: 2002-06-21
CURRENT APPLICATION NUMBER: US/10/179,131
NUMBER OF SEQ ID NOS: 10194
SEQ ID NO 3015
LENGTH: 1611
TYPE: DNA
ORGANISM: Candida albicans
US-10-179-131-3015

```

[illegible]

RESULT 12  
US-10-027-632-5967/c  
Sequence 5967, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108627.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23

```
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5967
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-5967
```

```
Query Match          2.0%; Score 40; DB 7; Length 610;
Best Local Similarity 55.1%; Pred. No. 0.6;
Matches 76; Conservative 1; Mismatches 61; Indels 0; Gaps 0;
```

```
QY 1845 AAGCTCAAGCTCATTCAGATGTTCACTTCAAGCTTCAAGCTCAAGCTCAAGCA 1904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 410 AGAGYGAACCTCACTCAAAATCTACTACTACTACTACTACTACTACTACTACTACT 351

QY 1905 CAATATAGTACGACATCCATCTTAACATATATAGCAACATCAATCAACCCCTG 1964
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 350 CTACTACTACTACTAATATATATATATATATATATATAGAGCTAAACATATTTCCCTG 291

QY 1965 ATCAACAATCAGATC 1982
    || ||| ||| |||
DB 290 TAGGCACTATCTCATC 273
```

```
RESULT 13
US-09-216-393B-340
; Sequence 340, Application US/09216393B
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 340
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(867)
; OTHER INFORMATION:
US-09-216-393B-340
```

```
Query Match          2.0%; Score 39.2; DB 5; Length 867;
Best Local Similarity 56.9%; Pred. No. 1.1;
Matches 91; Conservative 0; Mismatches 68; Indels 1; Gaps 1;
```

```
QY 1841 ACTGAAGTTCAGCTATCATCATGATAGTTCACACTTCAGCTCAGCTCAAC 1900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 436 actcaactcaacagcttcaccacagtcaggtacacagcagcagcagcagcagcagc 495

QY 1901 AGCACAATATATAGTACGATCCATCTTAACATATATAGCAACATCAATCAAC 1960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 496 actactaccactaccactaccactaccactaccactaccactaccactaccactacc 555

QY 1961 CCTGAT-CAACAAATCAGATCTCTCAACGACGACAACCA 1999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 556 acaaccacaactacacacacacacacacacacacacacacacacacacacacac 595
```

```
RESULT 14
US-09-216-393B-342
```

```
; Sequence 342, Application US/09216393B
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 342
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
US-09-216-393B-342
```

```
Query Match          2.0%; Score 39.2; DB 5; Length 867;
Best Local Similarity 56.9%; Pred. No. 1.1;
Matches 91; Conservative 0; Mismatches 68; Indels 1; Gaps 1;
```

```
QY 1841 ACTGAAGTTCAGCTATCATCATGATAGTTCACACTTCAGCTCAGCTCAAC 1900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 436 actcaactcaacagcttcaccacagtcaggtacacagcagcagcagcagcagcagc 495

QY 1901 AGCACAATATATAGTACGATCCATCTTAACATATATAGCAACATCAATCAAC 1960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 496 actactaccactaccactaccactaccactaccactaccactaccactaccactacc 555

QY 1961 CCTGAT-CAACAAATCAGATCTCTCAACGACGACAACCA 1999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 556 acaaccacaactacacacacacacacacacacacacacacacacacacacacac 595
```

```
RESULT 15
US-09-216-393B-343
; Sequence 343, Application US/09216393B
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 343
; LENGTH: 1397
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (238)..(1104)
; OTHER INFORMATION:
US-09-216-393B-343
```

```
Query Match          2.0%; Score 39.2; DB 5; Length 1397;
Best Local Similarity 56.9%; Pred. No. 1.4;
Matches 91; Conservative 0; Mismatches 68; Indels 1; Gaps 1;
```

```
QY 1841 ACTGAAGTTCAGCTATCATCATGATAGTTCACACTTCAGCTCAGCTCAAC 1900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 673 actcaactcaacagcttcaccacagtcaggtacacagcagcagcagcagcagcagc 732

QY 1901 AGCACAATATATAGTACGATCCATCTTAACATATATAGCAACATCAATCAAC 1960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 733 actactaccactaccactaccactaccactaccactaccactaccactaccactacc 792

QY 1961 CCTGAT-CAACAAATCAGATCTCTCAACGACGACAACCA 1999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 793 acaaccacacactacccaacaacgaagacaacca 832

Search completed: August 11, 2002, 04:23:38  
Job time: 8029 sec

---





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 19:00:18 ; Search time 2127.82 Seconds  
(without alignments)  
12679.838 Million cell updates/sec

Title: US-08-961-083-1  
Perfect score: 1999  
Sequence: 1 TAAATCTACGACATATAA.....ATCCTCACGACACACCA 1999

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues  
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estdb:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59.4	3.0	277	10	BM398655 5009-0-48
2	54.4	2.7	939	12	CNS00CNG
3	53.6	2.7	529	12	BM398931 AG-ND-121
4	52.8	2.6	447	9	AM502461 UI-HF-BRO
5	52.6	2.6	630	12	BM393736 AG-ND-136
6	51.8	2.6	691	12	BM384275 AG-ND-145
7	49.4	2.5	922	12	CNS0073M
8	49.2	2.5	1101	12	CNS00FXE
9	47.2	2.4	612	9	BE195101
10	47	2.4	934	9	BE195101 HVSMEH08
11	47	2.4	1749	10	BM415379 OP2165 M1
12	46.8	2.3	860	12	CNS018FL
13	46.6	2.3	674	12	BM399409 AG-ND-149
14	46.6	2.3	884	12	CNS006U0
15	46.4	2.3	751	12	BM381742 AG-ND-126
16	45.6	2.3	485	12	BM242511 AT2FC79TR
17	44.8	2.2	628	12	BM377839 AG-ND-167

18	44.4	2.2	419	9	AU033391	AU033391 AU033391
19	44.4	2.2	544	9	AU037837	AU037837 AU037837
20	44.4	2.2	895	12	CNS0071A	AL066286 Drosophila
21	43.8	2.2	988	12	CNS0072R	AL066743 Drosophila
22	43.6	2.2	475	10	BI594837	BI594837 AS-Lgc-44
23	43.4	2.2	472	10	BG114575	BG114575 602286235
24	43.4	2.2	582	12	AU060696	AU060696 AU060696
25	43	2.2	572	12	A2042338	A2042338 RPI-23-3
26	43	2.2	1101	12	CNS00LOO	AL066807 Drosophila
27	42.8	2.1	1043	12	AG135568	AG135568 Pan trogl
28	42.4	2.1	1101	12	CNS0100X	AL098379 Drosophila
29	42.2	2.1	532	12	A2556585	A2556585 RPI-23-2
30	42.2	2.1	544	12	A2800772	A2800772 2M059A03
31	42	2.1	606	12	BM396411	BM396411 AG-ND-139
32	42	2.1	998	12	CNS0144M	AL103696 Drosophila
33	41.4	2.1	579	10	C25758	C25758 C25758 Dict
34	41.2	2.1	1101	12	CNS0160F	AL107217 Drosophila
35	40.8	2.0	414	9	A1590574	A1590574 tw13h08.x
36	40.8	2.0	1072	12	CNS0070B	AL067418 Drosophila
37	40.8	2.0	1166	12	AG126229	AG126229 Pan trogl
38	40.6	2.0	690	12	B07870	B07870 XX37T7 PBu
39	40.6	2.0	829	9	AL569367	AL569367 AL569367
40	40.6	2.0	878	12	CNS0187R	AL108993 Drosophila
41	40.6	2.0	928	12	A2667900	A2667900 ENTG234TR
42	40.6	2.0	1101	12	CNS00LJT	AL068307 Drosophila
43	40.4	2.0	1036	12	CNS01ZIF	AL174336 Tetradon
44	40.4	2.0	1989	10	BF347552	BF347552 602020402
45	40.2	2.0	508	9	AU061871	AU061871 AU061871

#### ALIGNMENTS

RESULT 1  
BM398655/c 277 bp mRNA linear EST 17-JAN-2002  
LOCUS 5009-0-48-D02.c.1 Chilcoat/Turkewiltz cDNA (large fraction)  
DEFINITION Tetrahymena thermophila cDNA, mRNA sequence.  
ACCESSION BM398655.1 GI:18196708  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
EST.  
Tetrahymena thermophila.  
Tetrahymena thermophila  
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;  
Hymenostomatida; Tetrahymenina; Tetrahymena.

REFERENCE  
AUTHORS  
1 (bases 1 to 277)  
Turkewiltz,A.P., Karrer,K.M., Jahn,C., Orlas,E., Kirk,K.E., Frankel  
,J. and Klobutcher,L.  
EST from Tetrahymena thermophila, strain CU428.1, growing cells  
unpublished (2002)  
TITLE  
JOURNAL  
COMMENT  
Contact: Turkewiltz AP  
Molecular Genetics and Cell Biology  
University of Chicago  
920 E. 58th Street, Chicago, IL 60637, USA  
Tel: 773 702 4374  
Fax: 773 702 3172  
Email: apturkew@midway.uchicago.edu  
Seq primer: 73.

#### FEATURES

source  
Location/Qualifiers  
1..277  
/organism="Tetrahymena thermophila"  
/strain="CU428.1"  
/db\_xref="taxon:5911"  
/clone\_lib="Chilcoat/Turkewiltz cDNA (large fraction)"  
/note="Vector: Bluescript2 SK+, Details on library  
preparation can be found in Chilcoat and Turkewiltz (2001)  
Proc. Natl. Acad. Sci USA, 98: 8709-8713."

BASE COUNT  
ORIGIN  
51 a 69 c 99 g 58 t

Query Match 3.0%; Score 59.4; DB 10; Length 277;  
Best Local Similarity 54.9%; Pred. No. 0.00025;

[illegible]

RESULT	2
CNS00CNG	
LOCUS	939 bp DNA linear
DEFINITION	Drosophila melanogaster genome survey sequence TE13 end of BAC #
	BAR26H16 of RPc1-98 library from Drosophila melanogaster (fruit
	fly), genomic survey sequence.
ACCESSION	AL059400
VERSION	AL059400.1
KEYWORDS	GI:4946564
SOURCE	GSS.
ORGANISM	fruit fly.
	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 939)

REFERENCE 1 (bases 1 to 939)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

**COMMENT**

Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr  
- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr))  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org/TheBDGP/Drosophila>  
melanogaster BAC library was prepared by Kazutoyo Oosawa and  
Aaron Mamoser in Pieret de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPc1-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain y<sup>1</sup>; cn bw sp. the same strain used for the BDGP's  
p1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
vectors for hybridization from the BAC/PC Resource Center can be  
found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES	Location/Qualifiers
source	1. .939

BASE COUNT	ORIGIN
71 a	349 c 104 g 180 t 235 others

Query Match	2.7%	Score 54.4	DB 12,	Length 939;
Best Local Similarity	10.4%	Pred. No. 0.0076		
Best Match	21;	Conservative 122;	Mismatches 59;	Indels 0;
				Gaps 0;

Qy	1798	TCGCTACGTCGGAACACACCCCTCTCCACACAAACCCCCCATCACTGAAGTCAAGCTC	1857
Db	351	TTCCCTCAACATATTTTATGCCCCCCCCCTACACCTCTCCACACAAACCCCCCATCACTGAAGTCAAGCTC	410

[illegible]

RESULT	3			
LOCUS	BH398931/c			
DEFINITION	BH398931	529 bp	DNA	linear
	AG-ND-12J35..TF	ND-TAM	Anopheles	gambiae genomic clone AG-ND-12J35,
	DNA sequence.			

ACCESSION	BH398931
VERSION	BH398931.1
KEYWORDS	GI:17345147
SOURCE	GSS.
ORGANISM	African malaria mosquito.
REFERENCE	1. <i>Plasmodium falciparum</i> (3D7) whole-genome sequence.

ORGANISM  
Anopheles gambiae  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae  
; Anophelinae.  
1 (bases 1 to 529)  
REFERENCE

AUTHORSHIP: 2 (borders, 0.5)  
 AUTHORS: Shetty, J., Molek, J., Koo, H., Collins, F., Gardner, M. and Loftus, B. J.  
 TITLE: Direct Submission of BAC-end sequences from Anopheles gambiae  
 JOURNAL: Unpublished (2001)  
 COMMENT: Other\_GSSs: AG-ND-121J5.TR

Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 201 839 8200

Tel.: 301 838 0206  
 Fax: 301 838 3543  
 Email: bjloftus@sligr.org  
 This clone is from an A. gambiae BAC library (ND-TAM) developed by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.  
 Seq primer: M13 For  
 Class: BAC ends.

FEATURES	source
location/Qualifiers	1. .529
/organism="Anopheles gambiae"	
/strain="PEST"	
/db_xref="taxon:7165"	
/clone="AG-ND-121J5"	
/clone_1lb="ND-TAM"	
/note="Vector: pECBAC1; Site_1: HindIII"	
BASE COUNT	94 a 154 c 186 g 95 t
ORIGIN	

Query Match	2.7%	Score 53.6	DB 12	Length 529
Best Local Similarity	53.9%	Pred. No. 0.0096		
Matches 110: Conservative	0	Mismatches 94	Indels 0	Gaps 0

QY 26 CTCATTCGTGACTTGGGTCTCTCAAGCGCGGTCATGATCCCCAAGCTATGATATTTCCACA 85  
|||  
Db 278 CTCTGGCGGGAATTCGGCGAGAGCGCGCACCCCTGACCCCCCATCAAGAGATATCCCCAAG 219

Qy 86 GATTGGTAAAGCAATCGTTCTATGGAAGCACCATGGCTTTTGACACACAGGGGAT 145  
| | | | | | | | | | | | | | | |  
Db 218 GTCATGCAGGATGCGGTGCTGCCATCAGGACAACCCGCTTTTCGAGCACGTGGCTG 159

OY 146 GATACCATCGTATCCCTGGAGCTTTCTTGCCGCAATCTGCAGCAATTCCTCCAGGT 205  
 DB 158 GACTACAGAGGAGTGTGGCTGGCTGGCCACCTGGGCAAGGTCAAGACCCAGGG 99  
 OY 206 GATCAACTCTCACCCACAGTTG 229  
 DB 98 GCCTCCACATCACATGACAGTGTG 75

RESULT 4  
 AM502461/c 447 bp mRNA linear EST 01-MAR-2000  
 LOCUS UI-HF-BR0P-ajv-b-08-0-UI.r1 NIH\_MGC\_52 Homo sapiens cDNA clone  
 DEFINITION IMAGE:3075783 5', mRNA sequence.  
 ACCESSION AM502461  
 VERSION AM502461.1 GI:7116965  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 447)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Eco RI site shown at the beginning of the sequence.  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: M.B. Soares Lab  
 cDNA Library Arrayed by: M.B. Soares Lab  
 DNA Sequencing by: M.B. Soares Lab  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/dbp/image/image.html](http://www.bio.llnl.gov/dbp/image/image.html)  
 Seq primer: M13 Forward.

## FEATURES

source 1. 447  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3075783"  
 /clone\_1ib="NIH\_MGC\_52"  
 /tissue\_type="Lymph"  
 /cell\_type="germinal center B cells"  
 /cell\_line="MGC85"  
 /lab\_host="DH10B (LTI)"  
 /note="Vector: p7773-Pac; Site\_1: NotI; Site\_2: Eco RI;  
 Constructed from size fractionated cytoplasmic mRNA  
 (7.4-9.5kb). Directionally cloned. Cells provided by  
 Louis M. Staudt, Ph.D. Library preparation by Maria de  
 Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."  
 BASE COUNT 73 a 134 c 156 g 83 t 1 others  
 ORIGIN

Query Match 2.6%; Score 52.8; DB 9; Length 447;  
 Best Local Similarity 53.1%; Pred. No. 0.014;  
 Matches 11; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

OY 329 AAGCAAGAAATCTTGACTACTATATTAATAGGTACATGCTATGGAGACTATGGA 388  
 DB 447 AAGCGGAGATCTCTACCCGCTACCTGMACTGTCTGTCGGAAACAACATGTCGGC 388  
 OY 389 ATGCAGACAGACGCTCAAACTACTATGTAAAGACCTCAATTAATTAACTTACCTCAG 448  
 DB 387 ATCCAGAGAGCTGCCGAGCTACTTCGGGGTCAACGACATGCACTTGAACCTGGCAGCAG 328  
 OY 449 TTAGCCTTCTGGCTGGATGCTCAGGACCAACCAATATACCCCTATTACATCCA 508  
 DB 327 GGGGCTTCTGGCGCGGACATGTCGACGACGCGCGCTCAACCCNTACACCAACCC 268  
 OY 509 GAAGCAGCCCAAGACCGCAAACTGTGT 537

DB 267 GAGGGCGCCTGGCCGCGGCAACTGTGT 239

RESULT 5  
 BH393736 630 bp DNA linear GSS 11-DEC-2001  
 LOCUS AG-ND-136D6.TF ND-TAM Anopheles gambiae genomic clone AG-ND-136D6,  
 DEFINITION DNA sequence.  
 ACCESSION BH393736  
 VERSION BH393736.1 GI:17339877  
 KEYWORDS GSS.  
 SOURCE African malaria mosquito.  
 ORGANISM Anopheles gambiae  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae  
 ; Anophelinae.  
 REFERENCE 1 (bases 1 to 630)  
 AUTHORS Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.  
 TITLE Direct Submission of BAC-end sequences from Anopheles gambiae  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: [bjloftus@tigr.org](mailto:bjloftus@tigr.org)  
 This clone is from an A. gambiae BAC library (ND-TAM) provided by  
 F.H. Collins and sequenced by The Institute for Genomic Research  
 (TIGR). The BAC library was generated from A. gambiae PEST strain  
 DNA. All DNA was extracted from newly hatched first instar larvae  
 to minimize the inclusion of DNA from microorganisms that inhabit  
 the gut. The DNA is derived from mixed sexes of larvae. The BAC  
 library was constructed at Texas A&M University BAC Center  
 University, College Station, Texas 77843-2123, USA using a HindIII  
 partial digest.  
 Seq primer: M13 For  
 Class: BAC ends.

## FEATURES

source 1. 630  
 Location/Qualifiers  
 /organism="Anopheles gambiae"  
 /strain="PEST"  
 /db\_xref="taxon:7165"  
 /clone="AG-ND-136D6"  
 /clone\_1ib="ND-TAM"  
 /note="Vector: pECBAC1; Site\_1: HindIII"  
 BASE COUNT 120 a 200 c 189 g 121 t  
 ORIGIN

Query Match 2.6%; Score 52.6; DB 12; Length 630;  
 Best Local Similarity 64.2%; Pred. No. 0.018;  
 Matches 79; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

OY 75 ATATCCACAGATTTGGTTAAGCAATGTTCTATGCAAGACATCGCTTCTGAC 134  
 DB 504 AGATTTGCGCCCGCTGGTGAAGCACTGTCGACACCGAAGATCACGGTTCTACGAGC 563  
 OY 135 ACAGGGGATTGATTACATCCGTATCCTGGAGCTTCTTGGCGCAATCTGCAAGCAATT 194  
 DB 564 ACCAGCGATGACCTTCACCGCAGCCGTGGGCTGTGCATCATTAGCGCTGGCGCAATT 623  
 OY 195 CCC 197  
 DB 624 CCC 626

RESULT 6  
 BH384275 691 bp DNA linear GSS 10-DEC-2001  
 LOCUS AG-ND-145F13.TR ND-TAM Anopheles gambiae genomic clone AG-ND-145F13  
 DEFINITION DNA sequence.  
 ACCESSION BH384275





```

XX 13-SEP-2001 (Rel. 69, Created)
DT 13-SEP-2001 (Rel. 69, last updated, Version 1)
XX
DE OP165 Mixed Stage EST's from Globodera pallida, the potato cyst nematode
DE Globodera pallida cDNA, mRNA sequence.
XX
KM EST.
XX
OS Globodera pallida
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.
XX
RN 1-934
RP 1-934
RA Heer J., Sosinski B., Pokrzywa R.M., Warry A., Opperman C.;
RT "Mixed Stage EST's from Globodera pallida, the potato cyst nematode";
RL Unpublished.
XX
CC Contact: Opperman, C
CC Center for the Biology of Nematode Parasitism
CC NC State University; IACR-Rothamsted
CC Campus Box 7616; Raleigh, NC 27695, USA
CC Tel: 919.515.9500
CC Fax: 919.515.9500
CC Email: warthog@unity.ncsu.edu
XX No homology found. ; gtl1_2pcn_r_F10_pcn_r_079.abi.seq.screen.
XX
FH Key Location/Qualifiers
FH
FT source
FT 1. 934
FT /db_xref="taxon:36090"
FT /note="Vector: lambda GT11; This is a collaborative effort
FT between IACR-Rothamsted and North Carolina State
FT University. The library was constructed from mixed stage G.
FT pallida in lambda GT11 by Paul Burroughs,
FT IACR-Rothamsted.
FT /organism="Globodera pallida"
FT /clone_lib="Mixed Stage EST's from Globodera pallida, the
FT potato cyst nematode"
FT
XX
SQ Sequence 934 BP; 513 A; 262 C; 27 G; 132 T; 0 other;

```

```

Query Match      2.4%; Score 47; DB 3; Length 934;
Best Local Similarity 54.3%; Pred. No. 0.55;
Matches 95; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```

```

OY 1825 ACAACACCCCCCACTGAAAGTTCATCATCATGATGTTCACTTCACTGACAGTC 1884
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
DB 475 ACAAAACAAACAAACAAACAAACCAACCAACCAACCAACCAACCAACCA 534
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
OY 1885 TAGCTCAACCGCTCCAGCAAAATATAGTACGACTCAATCTTACATTAATACGCA 1944
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
DB 535 CACACCCCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 594
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
OY 1945 ACAATCAATCAACCCCTGATCAACAAATTCAGATCTCAACGACGACCAACCA 1999
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
DB 595 AAAAAAAAAAAAAACCCCACTCAACCAACACCCCAACCAACCAACCAACCA 649
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -

```

```

RESULT 11
LOCUS BM415379 1749 bp mRNA linear EST 28-JAN-2002
DEFINITION OP20453 Mixed Stage EST's from Globodera pallida, the potato cyst
nematode Globodera pallida cDNA, mRNA sequence.
ACCESSION BM415379
VERSION BM415379.1 GI:18381989
KEYWORDS EST.
SOURCE Globodera pallida.
ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.

```

```

REFERENCE 1 (bases 1 to 1749)
AUTHORS Heer J., Sosinski B., Pokrzywa R.M., Warry A. and Opperman C.
JOURNAL Mixed Stage EST's from Globodera pallida, the potato cyst nematode
Unpublished (2001)
COMMENT Contact: Opperman, C
Center for the Biology of Nematode Parasitism
NC State University; IACR-Rothamsted
Campus Box 7616; Raleigh, NC 27695, USA
Tel: 919.515.6699
Fax: 919.515.9500
Email: warthog@unity.ncsu.edu
gtl1_2pcn_r_F10_pcn_r_079.abi
Location/Qualifiers
FEATURES
source
1. 1749
/organism="Globodera pallida"
/db_xref="taxon:36090"
/clone_lib="Mixed Stage EST's from Globodera pallida, the
potato cyst nematode"
/note="Vector: lambda GT11; This is a collaborative effort
between IACR-Rothamsted and North Carolina State
University. The library was constructed from mixed stage
G. pallida in lambda GT11 by Paul Burroughs,
IACR-Rothamsted"
BASE COUNT 957 a 479 c 53 g 260 t
ORIGIN

```

```

Query Match      2.4%; Score 47; DB 10; Length 1749;
Best Local Similarity 54.3%; Pred. No. 0.72;
Matches 95; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```

```

OY 1825 ACAACACCCCCCACTGAAAGTTCATCATCATGATGTTCACTTCACTGACAGTC 1884
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
DB 475 ACAAAACAAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCA 534
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
OY 1885 TAGCTCAACCGCTCCAGCAAAATATAGTACGACTCAATCTTACATTAATACGCA 1944
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
DB 535 CACACCCCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 594
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
OY 1945 ACAATCAATCAACCCCTGATCAACAAATTCAGATCTCAACGACGACCAACCA 1999
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
DB 595 AAAAAAAAAAAAAACCCCACTCAACCAACACCCCAACCAACCAACCAACCA 649
    |||||  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -

```

```

RESULT 12
LOCUS CNS018FL 860 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACN13D02 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL109275
VERSION AL109275.1 GI:5629579
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 860)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billard at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
peloBAC11.
Location/Qualifiers
FEATURES

```

```

source
1. .860
/organism="Drosophila melanogaster"
/plasmid="pBelosBAC1"
/db.xref="taxon:7227"
/clone_11b="DrosBAC"
/clone="BACN13D02"
/note="end : 17"
218 a 134 c 113 g 139 t 256 others
BASE COUNT
ORIGIN

```

Query Match	2.3%	Score 46.8	DB 12	Length 860
Best Local Similarity	28.1%	Pred. No. 0.6		
Matches 78	Conservative 68	Mismatches 132	Indels 0	Gaps 0

[illegible]

RESULT	13
BH399409	
LOCUS	674 bp DNA linear GSS 11-DEC-2001
DEFINITION	AG-ND-149L20.TR ND-TAM Anopheles gambiae genomic clone AG-ND-149L20
ACCESSION	BH399409
VERSION	BH399409
KEYWORDS	BH399409.1 GI:17345625
SOURCE	GSS.
ORGANISM	African malaria mosquito.
	Anopheles gambiae
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
	Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae
	; Anophelinae.
REFERENCE	1 (bases 1 to 674)
AUTHORS	Sheely,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.
TITLE	Direct Submission of BAC-end sequences from Anopheles gambiae
JOURNAL	Unpublished (2001)
COMMENT	Other_GSSs: AG-ND-149L20..TF

Email: [bjofu@uic.edu](mailto:bjofu@uic.edu)  
This clone is from an *A. gambiae* BAC library (ND-74M) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from *A. gambiae* PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.  
Seq primer: M13 Rev  
Class: BAC ends.  
Location/Qualifiers

```

source      1..674
/organism="Anopheles gambiae"
/strain="PEST"
/db_xref="taxon:7165"
/clone="AG-ND-149420"
/clone_11b="ND-TAM"
/note="vector: PESTAC1; Site_1: HindIII"
BASE COUNT      129 a      202 g      124 t
ORIGIN

```

Query Match	2.3%	Score	46.6	DB	12	Length	674
Best Local Similarity	60.68%	Pred. NO.	0.61				
Matches	94	Conservative	0	Mismatches	59	Indels	2
						Gaps	1

[illegible]

RESULT	14
CNS00600/c	
LOCUS	CNS00600
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC #
LOCUS	BAC181N21 of RPCT-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL065923
VERSION	AL065923.1 GI:4944891
KEYWORDS	GSS.
SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 884)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a
COMMENT	

melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library> was prepared by Kazutoyo Osoegawa and Aaron Mermosser in Pieter de Jong's Laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

```

FEATURES
source      location/Qualifiers
            1. .884
              /organism="Drosophila melanogaster"
              /db_xref="taxon:7227"
              /clone_lib="RRC1-98"
              /clone="BAORI4N21"
              /note="end : 17"
230 a      62 c      139 g      124 t      329 others
ORIGIN

```

	Query Match	2.3%	Score 46.6;	DB 12;	Length 884;
	Best Local Similarity	10.6%;	Pred. No. 0.68;		
	Matches	20;	Conservative 106;	Mismatches 62;	Indels 0; Gaps 0;
OY	1811 AACCTACCTGCTGCACAACAACCCCATTCAGTGAAGTTCAAGSTCATCATCAGATAGT	1870			
	:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
Db	863 MYNN	804			
OY	1871 TCAACTTCACAGTGTAAGTCAACCACTCAACACACAATAATAGTAGCTACCAANTCCT	1930			
	:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
Db	803 MNN	744			
OY	1931 AACAATTAATAGCAACAATCAAAATCAACCCTGTATCAACAACAATCAGAATCCTCAACA	1990			
	:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
Db	743 MNMCMCMNMCMCMCMNMNMACACMACACMCACCMNMNMACSMNMNMCMCMACNMNMNM	684			
OY	1991 GCACAACC 1998				
	:: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
Db	683 MNNAAMNC 676				

RESULT	15
LOCUS	BH381742/c
DEFINITION	BH381742 751 bp DNA linear GSS 10-DEC-2001
ACCESSION	AG-ND-126M17.TR ND-TAM Anopheles gambiae genomic clone AG-ND-126M17
VERSION	'DNA sequence.'
KEYWORDS	BH381742
SOURCE	BH381742.1 GI:17327884
ORGANISM	GSS.
REFERENCE	African malaria mosquito.
AUTHORS	Anopheles gambiae
TITLE	Eukaryota: Metazoa; Arthropoda: Tracheata; Hexapoda: Insecta;
JOURNAL	Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae
COMMENT	; Anophel.
	1 (bases 1 to 751)
	Shetty,J., Malek,J., Koo,H., Collins,F., Gardiner,M. and Loftus,B.J.
	Direct Submission of BAC-end sequences from Anopheles gambiae
	Unpublished (2001)
	Other_GSSs: AG-ND-126M17.TR

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel.: 301 838 0208  
Fax: 301 838 3543  
Email: [bjloftus@tigr.org](mailto:bjloftus@tigr.org)  
This clone is from an *A. gambiæ* BAC library (ND-7AM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from *A. gambiæ* PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.  
Seq primer: M13 Rev  
Class: BAC end.

FEATURES	SOURCE	LOCATION/Qualifiers
	1..751	
	/organism="Anopheles gambiae"	
	/strain="PEST"	
	/db_xref="taxon:7165"	
	/clone="AG-ND-126M17"	
	/clone_lib="ND-TAM"	
	/note="vector: pECBAC1; site_1: HindIII"	
BASE COUNT	213 a 159 c 128 g 251 t	
ORIGIN		

Query Match	2.38;	Score 46.4;	DB 12;	Length 751;
Best Local Similarity	50.08;	Pred. No. 0.72;		
Matches 116; Conservative	0;	Mismatches 116;	Indels 0;	Gaps 0;

QY	266	CAGACTATTTCGTAAGGGCTCAGAAAGCTGGTTTACGAGATCGATTGAAACAAAAGCA	335
QY	266		
Db	251	CAAAAGGATTTCAGAAAGCTTAAAGAAATGGTCTGTACGGGTAGCCTGGAGAAAGAGGTAT	132
QY	326	ACCAAGCAGAAGAACTCTTGACCTACTATATATAATTAAGGCTCTACATGTCTAAATGGGAACAT	385
Db	191	ACGAAAGAAAGAAATCATATACACTGTATTTTCAATTAATTCGATTTTACCTCAATGCAAT	132
QY	366	GGATTCGACGACGACGCTCAAAACTACTATGTGTAAAGCCTCAATATTTTAAGTTTACCT	445
QY	366		
Db	131	GGAAATTTGAAATGGGCTTCAGGAATCTATTTCAATTAACATACGAATGGAGCTTACCCCTCT	72
QY	446	CAGTTAGCCTTGCGGCTGGAATCCCTAGGCACCAACCAATATGAGCCCT	497
QY	446		
Db	71	GAACTCGAATGTTGTGGCATCTCTGGAAAGCTCCGAATGCAACAACCCCT	20

Search completed: August 11, 2002, 02:09:42  
Job time: 25764 sec



---





QY 61 TGCCCAAGCTATGATATATTCACAGATTGGTTAAGCAATCGTTTCTATCGAAGACCA 120  
 |||||  
 Db 61 TGCCCAAGCTATGATATATTCACAGATTGGTTAAGCAATCGTTTCTATCGAAGACCA 120  
 QY 121 TCGCTTCTTCGACCAAGAGGGGATGATACATCCGTAATCTCGGAGCTTTCTTCGCAA 180  
 |||||  
 Db 121 TCGCTTCTTCGACCAAGAGGGGATGATACATCCGTAATCTCGGAGCTTTCTTCGCAA 180  
 QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCCACCAAGCTGATTAAGTTGAC 240  
 |||||  
 Db 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCCACCAAGCTGATTAAGTTGAC 240  
 QY 241 TTACTTTTCAACTTCGACTTCGACAGACATATTTCTGTAAGGCTCAGGAAGCTTGT 300  
 |||||  
 Db 241 TTACTTTTCAACTTCGACTTCGACAGACATATTTCTGTAAGGCTCAGGAAGCTTGT 300  
 QY 301 AGCGATTGATTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATTAATTA 360  
 |||||  
 Db 301 AGCGATTGATTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATTAATTA 360  
 QY 361 GGTGACATGCTAATGGGAAGTATGAAATGAGACAGACAGCTCAAACTACTATGTTAA 420  
 |||||  
 Db 361 GGTGACATGCTAATGGGAAGTATGAAATGAGACAGACAGCTCAAACTACTATGTTAA 420  
 QY 421 AGACCTCAATATTTAAGTTTACCTCACTGATGCTTGGCTGGGAATGCTTCAGAGCAC 480  
 |||||  
 Db 421 AGACCTCAATATTTAAGTTTACCTCACTGATGCTTGGCTGGGAATGCTTCAGAGCAC 480  
 QY 481 AAACCAATATACCCCTTATTCATCCAGAAAGAGCCCAAGAGCCGCAAACTGGTCTT 540  
 |||||  
 Db 481 AAACCAATATACCCCTTATTCATCCAGAAAGAGCCCAAGAGCCGCAAACTGGTCTT 540  
 QY 541 ATCTGAATATGAAAAATCAAGGCTACATCTGCTGAAACAGTATGAAAGAGAGTCAATAC 600  
 |||||  
 Db 541 ATCTGAATATGAAAAATCAAGGCTACATCTGCTGAAACAGTATGAAAGAGAGTCAATAC 600  
 QY 601 ACCAATATCTGATGAGTACAAAGTCTCAAAATCAGCAAGTAAATACCTGCTTACATGGA 660  
 |||||  
 Db 601 ACCAATATCTGATGAGTACAAAGTCTCAAAATCAGCAAGTAAATACCTGCTTACATGGA 660  
 QY 661 TAAATTAATCTCAAGGAGTCAATCAATGTTGAGAGAGAAACAGGCTTAACTACTAC 720  
 |||||  
 Db 661 TAAATTAATCTCAAGGAGTCAATCAATGTTGAGAGAGAAACAGGCTTAACTACTAC 720  
 QY 721 AACTGGATGAGTGTCTACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780  
 |||||  
 Db 721 AACTGGATGAGTGTCTACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780  
 QY 781 TTACATATACAGAGCAATACGTTGCCATTCAGAGAGTGAATTTGCAAGTCCGCTTACCAT 840  
 |||||  
 Db 781 TTACATATACAGAGCAATACGTTGCCATTCAGAGAGTGAATTTGCAAGTCCGCTTACCAT 840  
 QY 841 TGTGATGTTTCTAAGCTTAAGTCAATGCTCCAGCTAGGAGAGCCCATTCAGTCAAGTAA 900  
 |||||  
 Db 841 TGTGATGTTTCTAAGCTTAAGTCAATGCTCCAGCTAGGAGAGCCCATTCAGTCAAGTAA 900  
 QY 901 TGTTCCTTGGGAATTAACAGAGTGAAGAAACAACCGGAGCTGGGGATCAACATATGAA 960  
 |||||  
 Db 901 TGTTCCTTGGGAATTAACAGAGTGAAGAAACAACCGGAGCTGGGGATCAACATATGAA 960  
 QY 961 ACCGATCAAGACTATGCTCCTGCTTGGAGTAGGGTCTACAGATTCAGTCTACTAT 1020  
 |||||  
 Db 961 ACCGATCAAGACTATGCTCCTGCTTGGAGTAGGGTCTACAGATTCAGTCTACTAT 1020  
 QY 1021 CGTTACAGATAGCCCTTAATACCTCGGAGCAAAATCTCCTGTTTAACTGGGATAG 1080  
 |||||  
 Db 1021 CGTTACAGATAGCCCTTAATACCTCGGAGCAAAATCTCCTGTTTAACTGGGATAG 1080  
 QY 1081 GGGCTACTTGGCAACATCAGCTTGACATAGCCCTGCAACAAATCGCAAAAGCTCCAGC 1140  
 |||||  
 Db 1081 GGGCTACTTGGCAACATCAGCTTGACATAGCCCTGCAACAAATCGCAAAAGCTCCAGC 1140

QY 1141 CGTGGAAACTCTAACAAGGTGCGACTGACACCGGCCAAGACTTTCTTAAATGCTTAGG 1200  
 |||||  
 Db 1141 CGTGGAAACTCTAACAAGGTGCGACTGACACCGGCCAAGACTTTCTTAAATGCTTAGG 1200  
 QY 1201 AATGCACTACCAAGTATCTACTCTCAAAATGCCATTTCAAGTAAACAAACGGAATCGA 1260  
 |||||  
 Db 1201 AATGCACTACCAAGTATCTACTCTCAAAATGCCATTTCAAGTAAACAAACGGAATCGA 1260  
 QY 1261 CAAAAATATGAGAGCAATGTAAGAAAGATGGCTGCTTACGCTGCTTTGCAAAATGG 1320  
 |||||  
 Db 1261 CAAAAATATGAGAGCAATGTAAGAAAGATGGCTGCTTACGCTGCTTTGCAAAATGG 1320  
 QY 1321 TGAACCTTACTATAAACCAATGATATATCCATTAAGTCTTTTATGATGAGGAGTAA 1380  
 |||||  
 Db 1321 TGAACCTTACTATAAACCAATGATATATCCATTAAGTCTTTTATGATGAGGAGTAA 1380  
 QY 1381 AGAGTCTCTAATGTCGGAACCTCGGCCATGGAAGAAACGACAGCCTATATGATGACCGA 1440  
 |||||  
 Db 1381 AGAGTCTCTAATGTCGGAACCTCGGCCATGGAAGAAACGACAGCCTATATGATGACCGA 1440  
 QY 1441 CATGATGAAAAACAGTCTGACTTATGAACTGAGCAAGTATGCTTGGCTGCTCC 1500  
 |||||  
 Db 1441 CATGATGAAAAACAGTCTGACTTATGAACTGAGCAAGTATGCTTGGCTGCTCC 1500  
 QY 1501 TCAGGCTGTAAACAGGAACTCTAATATACAGACAGAGAAATGAAACCAATCAA 1560  
 |||||  
 Db 1501 TCAGGCTGTAAACAGGAACTCTAATATACAGACAGAGAAATGAAACCAATCAA 1560  
 QY 1561 GACCTCTAATTTGTAGACCTGATGAATTTTGTGCTGCTATACGCTTAATTTCAAT 1620  
 |||||  
 Db 1561 GACCTCTAATTTGTAGACCTGATGAATTTTGTGCTGCTATACGCTTAATTTCAAT 1620  
 QY 1621 GGTGATGAGAGGCTTTCTTAAACCGTCTGACACCACTTGTAGGCAATGGCTTACGCT 1680  
 |||||  
 Db 1621 GGTGATGAGAGGCTTTCTTAAACCGTCTGACACCACTTGTAGGCAATGGCTTACGCT 1680  
 QY 1681 CGTGGCAAGTTTACGCTTATGATGAGTACCTGCTGTAAGAGCAATTCAGAAAGA 1740  
 |||||  
 Db 1681 CGTGGCAAGTTTACGCTTATGATGAGTACCTGCTGTAAGAGCAATTCAGAAAGA 1740  
 QY 1741 TTGGAATATACCAAGGGGCTCTACAGAAATGGAATTCGTATTTAAATGGTGGCTG 1800  
 |||||  
 Db 1741 TTGGAATATACCAAGGGGCTCTACAGAAATGGAATTCGTATTTAAATGGTGGCTG 1800  
 QY 1801 TTCTACGTGGAAGTCACTGCTCCACAAACACCCCATCACTGAAGTTCAAGCTCATC 1860  
 |||||  
 Db 1801 TTCTACGTGGAAGTCACTGCTCCACAAACACCCCATCACTGAAGTTCAAGCTCATC 1860  
 QY 1861 ATCAGATATGTTCAACTTCACAGTCTAGCTCAACCACTTCAAGCAAAATTAATAGTACGAC 1920  
 |||||  
 Db 1861 ATCAGATATGTTCAACTTCACAGTCTAGCTCAACCACTTCAAGCAAAATTAATAGTACGAC 1920  
 QY 1921 TACCAATCTTAACATATATAGCAACATCAATTAACAAACCCCTGATCAACAAATTCAGAA 1980  
 |||||  
 Db 1921 TACCAATCTTAACATATATAGCAACATCAATTAACAAACCCCTGATCAACAAATTCAGAA 1980  
 QY 1981 TCTCAACACAGCAACCA 1999  
 |||||  
 Db 1981 TCTCAACACAGCAACCA 1999

RESULT 2  
 US-08-481-435-5  
 ; Sequence 5, Application US/08481435  
 ; Patent No. 6027906  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baljanes, Tanjore S  
 ; APPLICANT: Town, Christine  
 ; TITLE OF INVENTION: No. 6027906el Polypeptides  
 ; NUMBER OF SEQUENCES: 42  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: White & Case  
 ; STREET: 1155 Avenue of the Americas

CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/481,435  
APPLICATION NUMBER: US/08/481,435  
FILING DATE: 10-JUL-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IN 580/MAS/94  
FILING DATE: 01-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9404072-2  
FILING DATE: 24-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sterner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-151  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2049 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORGANISM: Streptococcus pneumoniae  
STRAIN: PM 1  
IMMEDIATE SOURCE:  
LIBRARY: PCR cloning  
CLONE: pARC 0512 Soluble PBP 1A del 38  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2049  
FEATURE:  
NAME/KEY: mat.peptide  
LOCATION: 1..2046  
US-08-481-435-5

Query Match 99.1%; Score 1981.4; DB 3; Length 2049;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTAGACATAAAATCAATCACTATTGCTGACTTGGGTTCTGAACGCCGGCTCAA 60  
DB 48 TAAATCTAGACATAAAATCAATCACTATTGCTGACTTGGGTTCTGAACGCCGGCTCAA 107  
QY 61 TGCCCACTATGATATTTCCACAGATTGGTTAAGCAATCGTTTCTATGTAAGACCA 120  
DB 108 TGGCCAACTATGATATTTCCACAGATTGGTTAAGCAATCGTTTCTATGTAAGACCA 167  
QY 121 TCGCTTCTGACACGAGGGGATGATACCTCCGATCCGAGGAGCTTTCTTCCGCA 180  
DB 168 TCGCTTCTGACACGAGGGGATGATACCTCCGATCCGAGGAGCTTTCTTCCGCA 227  
QY 181 TCTGCAAGCAATTCCTCCACAGGTGATCACTCCACCAAGTTGATTAACTGAC 240  
DB 228 TCTGCAAGCAATTCCTCCACAGGTGATCACTCCACCAAGTTGATTAACTGAC 287  
QY 241 TTACTTTCACTTCGACTTCGACACGACTATTTCTGTAAGGCTCGAAGAGCTTGGT 300  
DB 288 TTACTTTCACTTCGACTTCGACACGACTATTTCTGTAAGGCTCGAAGAGCTTGGT 347  
QY 301 AGGATTCAGTTAGAACAAAAGCAACGACAGAAATCTTGACCTATATATAATA 360

|||||  
DB 348 AGCGATTCAGTTAGAACAAAAGCAACGACAGAAATCTTGACCTATATATAATA 407  
QY 361 GGTCTACATGCTTAATGGGAATATGGAATGACAGACAGCTCAAACTACTATGTGA 420  
DB 408 GGTCTACATGCTTAATGGGAATATGGAATGACAGACAGCTCAAACTACTATGTGA 467  
QY 421 AGACCTCATATATTAAGTTACCTCAGTTAGCTTGGGCTGGAATGCCAGGACAC 480  
DB 468 AGACCTCATATATTAAGTTACCTCAGTTAGCTTGGGCTGGAATGCCAGGACAC 527  
QY 481 AAACCAATATGACCCCTATTTCATCCAGAACGACCCCAAGACCGCCGAACCTTGCTT 540  
DB 528 AAACCAATATGACCCCTATTTCATCCAGAACGACCCCAAGACCGCCGAACCTTGCTT 587  
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAATAC 600  
DB 588 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAATAC 647  
QY 601 ACCAATTCGATGGAGCTACAAAGTCTCAAAATCAGCAAGTATTAACCTGCTACATGA 660  
DB 648 ACCAATTCGATGGAGCTACAAAGTCTCAAAATCAGCAAGTATTAACCTGCTACATGA 707  
QY 661 TAATTACCTCAAGAGATCATCAATCAAGTTGAAGAAACAGGCTATTAACCTACTAC 720  
DB 708 TAATTACCTCAAGAGATCATCAATCAAGTTGAAGAAACAGGCTATTAACCTACTAC 767  
QY 721 AACTGGATGATGTCCTACCAAAATGTAGACCAGAAAGCTCAAAAACATCTGTGGAT 780  
DB 768 AACTGGATGATGTCCTACCAAAATGTAGACCAGAAAGCTCAAAAACATCTGTGGAT 827  
QY 781 TTACAATACAGCAATACGCTTACCTATCCAGCAGTGAATGAAGTCCGCTTACCAT 840  
DB 828 TTACAATACAGCAATACGCTTACCTATCCAGCAGTGAATGAAGTCCGCTTACCAT 887  
QY 841 TGTGATGCTTTTAACGGTAAGTCAATGCTCCAGCTAGAGACCCGATCACTCAAGTGA 900  
DB 888 TGTGATGCTTTTAACGGTAAGTCAATGCTCCAGCTAGAGACCCGATCACTCAAGTGA 947  
QY 901 TGTTCCTTGGGAATTAACCAAGTAGAAGAAACACCGCATGGGATCACTATGAA 960  
DB 948 TGTTCCTTGGGAATTAACCAAGTAGAAGAAACACCGCATGGGATCACTATGAA 1007  
QY 961 ACCGATACAGACTATGCTCCCTGCTGGAGTAGAGGTGCTACGATCAATCGTACTAT 1020  
DB 1008 ACCGATACAGACTATGCTCCCTGCTGGAGTAGAGGTGCTACGATCAATCGTACTAT 1067  
QY 1021 CGTTACGATGAGCCCTATTAATCTACCTGGGACAAATACCTGTTTAACTGGGATG 1080  
DB 1068 CGTTACGATGAGCCCTATTAATCTACCTGGGACAAATACCTGTTTAACTGGGATG 1127  
QY 1081 GGGCTACTTTGGCAACATCACTTGCCTGCAATACGCTGCAACATGCGGAACGCTCCAC 1140  
DB 1128 GGGCTACTTTGGCAACATCACTTGCCTGCAATACGCTGCAACATGCGGAACGCTCCAC 1187  
QY 1141 CGTGAACCTTAACCAAGGTGGAATCAACCGCGCAAACTTTCCTAAATGGTCTAG 1200  
DB 1188 CGTGAACCTTAACCAAGGTGGAATCAACCGCGCGCAAGACTTTCCTAAATGGTCTAG 1247  
QY 1201 AATGACTACCAATATTAATCACTACTCAATGCAATTCATTAATCAACACAAACGAATCAGA 1260  
DB 1248 AATGACTACCAATATTAATCACTACTCAATGCAATTCATTAATCAACACGAATCAGA 1307  
QY 1261 CAAAAAATATGAGCAAGTAGTAAGAAAGTGGCTGCTGTTAGCGCTTTGCAAAATG 1320  
DB 1308 CAAAAAATATGAGCAAGTAGTAAGAAAGTGGCTGCTGTTAGCGCTTTGCAAAATG 1367  
QY 1321 TGGAACTTACTATTAACCAATGATATATCATTAATGCTTATGATGGAGGAA 1380  
DB 1368 TGGAACTTACTATTAACCAATGATATATCATTAATGCTTATGATGGAGGAA 1427  
QY 1381 AGAGTTCCTAATGTCGGAACCTGCGCATGAAGAAACGACGCTATATGATGACGA 1440  
|||||

Db	1428	AGAGTTCCTAAATGTCGGAACCTGTGGCATGAAAGAAAGACAGCCCTAATGATGACGA	1487
Qy	1441	CATGATGAAAACAGCTCTTGACTTATGGAACGTGAGCAAAATGCCATCTTGGCTGCC	15000
Db	1488	CATGATGAAAACAGCTCTTGATGATGGAACGTGAGCAAAATGCCATCTTGGCTGCC	15478
Qy	1501	TCAGAGCTGGTAAACAGGAACCTCTAATATACAGAGGAGGAAATTGSAAACCCATCA	15668
Db	1548	TCAGAGCTGGTAAACAGGAACCTCTAATATACAGAGGAGGAAATTGSAAACCCATCA	1607
Qy	1561	GACCTTCGAATTTGTAGCACACCTGATGAACATATTGCTGGCTATAGCGCTAAATATTCAT	1620
Db	1608	GACCTTCGAATTTGTAGCACACCTGATGAACATATTGCTGGCTATAGCGCTAAATATTCAT	1667
Qy	1621	GGCTGTATGAGCAGAGCTATTCTAACCGCTGTGACACACATTTGAGCAATGGCTTACGT	16800
Db	1668	GGCTGTATGAGCAGAGCTATTCTAACCGCTGTGACACACATTTGAGCAATGGCTTACGT	17277
Qy	1681	CGCTGCCAAAGTTTACCGGCTCTATGATGACCTTCCTGTGGAAGGAGCAATCCGAGA	17400
Db	1728	CGCTGCCAAAGTTTACCGGCTCTATGATGACCTTCCTGTGGAAGGAGCAATCCGAGA	17878
Qy	1741	TTTGAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGTCG	18000
Db	1788	TTTGAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGTCG	18477
Qy	1801	TTTCTACGTGGGAACCTCACCTGCTCCACAAACACCCCATCATACTGAAAGTTCAGCTATC	18600
Db	1848	TTTCTACGTGGGAGCTCACCTGCTCCACAAACACCCCATCATACTGAAAGTTCAGCTATC	19077
Qy	1861	ATCAGATAGTTCAACTCAGACGTCGTAGCTCAACCATCCACACAAATTAATAGTACAC	19200
Db	1908	ATCAGATAGTTCAACTCAGACGTCGTAGCTCAACCATCCACACAAATTAATAGTACAC	19677
Qy	1921	TACCAATCTTACCAATATATACGCAACATCAATATACAAACCCCTGATCAACAAATACAG	19800
Db	1968	TACCAATCTTACCAATATATACGCAACATCAATATACAAACCCCTGATCAACAAATACAG	20277
Qy	1981	TTCCCAACGACGCAACCA 1999	
Db	2028	TTCTTCAACGACGCAACCA 2046	

RESULT 3  
US-08-245-511-3  
Sequence 3, Application US/08245511  
Patent No. 5928900  
GENERAL INFORMATION:  
APPLICANT: Measure, H Robert  
APPLICANT: Pearce, Barbara J  
APPLICANT: Tuomanen, Elaine  
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND  
CELLULAR VACCINES BASED THEREON  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/245,511  
FILING DATE: 18-MAY-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/116,541  
FILING DATE: 01-SEP-1994

:	ATTORNEY/AGENT INFORMATION:
:	NAME: Jackson Esq., David A.
:	REGISTRATION NUMBER: 26,742
:	REFERENCE/DOCKET NUMBER: 600-1-069 CIP
:	TELECOMMUNICATION INFORMATION:
:	TELEPHONE: 201 487-5800
:	TELEFAX: 201 343-1684
:	TELEX: 133521
:	INFORMATION FOR SEQ ID NO: 3:
:	SEQUENCE CHARACTERISTICS:
:	LENGTH: 960 base pairs
:	TYPE: nucleic acid
:	STRANDEDNESS: both
:	TOPOLOGY: unknown
:	MOLECULE TYPE: DNA (genomic)
:	HYPOTHETICAL: NO
:	ANTI-SENSE: NO
:	ORIGINAL SOURCE:
:	ORGANISM: Streptococcus pneumoniae
:	STRAIN: R6
:	IMMEDIATE SOURCE:
:	CLONE: SPRU42
:	FEATURE:
:	NAME/KEY: CDS
:	LOCATION: 1..960
:	US-08-245-511-3
Query Match	47.2%; Score 944.2; DB 2; Length 960;
Best Local Similarity	99.7%; Pred. No. 1.9e-278;
Matches 946; Conservative	0; Mismatches 3; Indels 0; Gaps
QY	1 TAAATTCAGCACAATAAAAAATCAACTCATTTGCTGACTTGGTTCTGAACGCCGCCTCAA 60
DB	12 TAAATTCAGCACAATAAAAAATCAACTCATTTGCTGACTTGGTTCTGAACGCCGCCTCAA 71
QY	61 TGCCCAACTAATGATATATCCCACAGATTGTGTTAAGSCAATCGTTTCTATCGAAGCCA 120
DB	72 TGCCCAACTAATGATATATCCCACAGATTGTGTTAAGSCAATCGTTTCTATCGAAGCCA 131
QY	121 TCGCTTTCGACCACAGGGGGATTGATACCATCCGTAATCCCTGGAGCTTCTTGCGCAA 180
DB	132 TCGCTTTCGACCACAGGGGGATTGATACCATCCGTAATCCCTGGAGCTTCTTGCGCAA 191
QY	181 TCTGCAAGCAATTCCTCCACAGGTGATCTCAACTCACCTCACCCAACAGTTGATTAAGTGAC 240
DB	192 TCTGCAAGCAATTCCTCCACAGGTGATCTCAACTCACCTCACCCAACAGTTGATTAAGTGAC 251
QY	241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTTCTGTAAGGCTCAGAGAAGCTTGCTT 300
DB	252 TTACTTTTCAACTTCGACTTCGACACAGACTATTTTCTGTAAGGCTCAGAGAAGCTTGCTT 311
QY	301 AGCGATTGAGTTAGAACAAAAGACACCAACCAAGAAATCTTGACTATATATAATAA 360
DB	312 AGCGATTGAGTTAGAACAAAAGACACCAACCAAGAAATCTTGACTATATATAATAA 371
QY	361 GGTCTACATGCTCTAATGGAAGCACTATGCAATGCAGACAGACAGCTTCAAACTACTATGGTAA 420
DB	372 GGTCTACATGCTCTAATGGAAGCACTATGCAATGCAGACAGACAGCTTCAAACTACTATGGTAA 431
QY	421 AGACCTCAATATATTTAAATTTTACCTCAGTTAGCCTTCCTGGCTGGAATGCCTCAGGCACC 480
DB	432 AGACCTCAATATATTTAAATTTTACCTCAGTTAGCCTTCCTGGCTGGAATGCCTCAGGCACC 491
QY	481 AAACCATATATAGCCCTATTTACATCCAGAGACAGCCCAAGCCGCGGAAACTTGGTCTT 540
DB	492 AAACCATATATAGCCCTATTTACATCCAGAGACAGCCCAAGCCGCGGAAACTTGGTCTT 551
QY	541 ATCTGAATGAAAAATCAAGGCTCATCTCTGCGAACAGATGAGAAACAGTCAATAC 600
DB	552 ATCTGAATGAAAAATCAAGGCTCATCTCTGCGAACAGATGAGAAACAGTCAATAC 611
QY	601 ACCATTTCATATGAGACACAATCTCAANTCAGCAAGTAATTAACCTGCTTACATGGA 660

|||||  
Db 612 ACCAATTACTGATGGGCTCAAAAGTCTCAATACAGAACTAATTCACCTGCTTACTGGA 671  
Qy 661 TAAATACCTCAAGGAAGTATCATATCAATGTAAGAAAGAAAGGCTATACCTACTGAC 720  
Db 672 TAATTACCTCAAGGAAGTATCATATCAATGTAAGAAAGAAAGGCTATACCTACTGAC 731  
Qy 721 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780  
Db 732 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 791  
Qy 781 TTACATACAGACGATATAGCTTCCCTATCCAGACGATGATGTAATGCAAGTCCCTTACCAT 840  
Db 792 TTACATACAGACGATATAGCTTCCCTATCCAGACGATGATGTAATGCAAGTCCCTTACCAT 851  
Qy 841 TGTGATGTTTAAACGGTAAGTCAATGCTCCAGCTAGAGACAGCCATCACTCAAGTAA 900  
Db 852 TGTGATGTTTAAACGGTAAGTCAATGCTCCAGCTAGAGACAGCCATCACTCAAGTAA 911  
Qy 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGCTGGGGA 949  
Db 912 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGCTGGGGA 960

RESULT 4  
US-08-600-993A-3  
; Sequence 3, Application US/08600993A  
; Patent No. 5981229  
; GENERAL INFORMATION:  
; APPLICANT: Masure, H Robert  
; APPLICANT: Pearce, Barbara J  
; APPLICANT: Tuomanen, Elaine  
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/600,993A  
; FILING DATE: 1-MAR-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/116,541  
; FILING DATE: 01-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-069 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 960 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pneumoniae  
; STRAIN: R6  
; IMMEDIATE SOURCE:  
; CLONE: SPRU42  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..960  
; US-08-600-993A-3

Query Match 47.2%; Score 944.2; DB 2; Length 960;  
Best Local Similarity 99.7%; Pred. No. 1.9e-278;  
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 TAAATCTACGACAAATAAATCACTCATTTGCTGACTTGGTTCGAACGCCGCTCAA 60  
Db 12 TAAATCTACGACAAATAAATCACTCATTTGCTGACTTGGTTCGAACGCCGCTCAA 71  
Qy 61 TGCCCAAGCTAATGATATTTCCACAGATTTGTTAAGCAATGCTTATGGAAGCA 120  
Db 72 TGCCCAAGCTAATGATATTTCCACAGATTTGTTAAGCAATGCTTATGGAAGCA 131  
Qy 121 TCGCTTCTCGACACAGAGGGGATTTGATACACCTATCTGGAGCTTTCTGGGCAA 180  
Db 132 TCGCTTCTCGACACAGAGGGGATTTGATACACCTATCTGGAGCTTTCTGGGCAA 191  
Qy 181 TCTGCAAGCAATTTCCCTCAAGGTGATCACTCCACCAACAGTTGATTAAGTTGAC 240  
Db 192 TCTGCAAGCAATTTCCCTCAAGGTGATCACTCCACCAACAGTTGATTAAGTTGAC 251  
Qy 241 TTAATTTCACTTCGACTTCGACCAAGCTATTTCTGTTAAGCTCAGAGAGCTTGGT 300  
Db 252 TTAATTTCACTTCGACTTCGACCAAGCTATTTCTGTTAAGCTCAGAGAGCTTGGT 311  
Qy 301 AGCGATTCACTTAGAACAACCAAGCAAGCAAGAAATCTTGACTACTATTAATAA 360  
Db 312 AGCGATTCACTTAGAACAACCAAGCAAGCAAGAAATCTTGACTACTATTAATAA 371  
Qy 361 GGTCTACATGCTAATGGGAACATGATGACAGACAGCAGCTCAAACTACTATGTTAA 420  
Db 372 GGTCTACATGCTAATGGGAACATGATGACAGACAGCAGCTCAAACTACTATGTTAA 431  
Qy 421 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCTTGTGCTGTAAGTGCCTCAGCACC 480  
Db 432 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCTTGTGCTGTAAGTGCCTCAGCACC 491  
Qy 481 AAACCAATATGACCCCTATTCACATCCAGAAGACCCCAAGACCCGCCAAACTTGGTCT 540  
Db 492 AAACCAATATGACCCCTATTCACATCCAGAAGACCCCAAGACCCGCCAAACTTGGTCT 551  
Qy 541 ATCTGAAATGAAAAATCAAGGGCTACATCTCTGCTGAACAGTATGAAGAACACTCAATAC 600  
Db 552 ATCTGAAATGAAAAATCAAGGGCTACATCTCTGCTGAACAGTATGAAGAACACTCAATAC 611  
Qy 601 ACCAATTACTGATGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGA 660  
Db 612 ACCAATTACTGATGATGCTCAAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGA 671  
Qy 661 TAAATACCTCAAGGAAGTATCATATCAATGTAAGAAAGAAAGGCTATACCTACTGAC 720  
Db 672 TAAATACCTCAAGGAAGTATCATATCAATGTAAGAAAGAAAGGCTATACCTACTGAC 731  
Qy 721 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780  
Db 732 AACTGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 791  
Qy 781 TTACATACAGACGATATAGCTTCCCTATCCAGACGATGATGTAATGCAAGTCCCTTACCAT 840  
Db 792 TTACATACAGACGATATAGCTTCCCTATCCAGACGATGATGTAATGCAAGTCCCTTACCAT 851

OY 841 TGTGTATGTTCTTAACGAGTAAAGTCATGTCAGTAGAGACGCCATCATGATCAAGTAA 900  
 |||||  
 DB 852 TGTGTATGTTCTTAACGAGTAAAGTCATGTCAGTAGAGACGCCATCATGATCAAGTAA 911  
 OY 901 TGTTCCTTGGGAATTAACCAAGCAGTAGAAACCAACCGCAGCTGGGA 949  
 |||||  
 DB 912 TGTTCCTTGGGAATTAACCAAGCAGTAGAAACCAACCGCAGCTGGGA 960

RESULT 5

US-08-743-637B-27/c  
 ; Sequence 27, Application US/08743637B  
 ; Patent No. 5994066  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BERGERON, Michel G.  
 ; APPLICANT: PICARD, Francois J.  
 ; APPLICANT: OUELLETTE, Marc  
 ; APPLICANT: ROY, Paul H.  
 ; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA  
 ; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND  
 ; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED  
 ; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...  
 ; NUMBER OF SEQUENCES: 273  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: O'ARLES & BRADY  
 ; STREET: 411 EAST WISCONSIN AVENUE  
 ; CITY: MILWAUKEE  
 ; STATE: WISCONSIN  
 ; COUNTRY: USA  
 ; ZIP: 53202-4497  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/743,637B  
 ; FILING DATE: 04-NOV-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/526,840  
 ; FILING DATE: 11-SEP-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BAKER, Jean C.  
 ; REGISTRATION NUMBER: 35,433  
 ; REFERENCE/DOCKET NUMBER: 850586.90012  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (414) 277-5000  
 ; TELEFAX: (414) 277-5591  
 ; INFORMATION FOR SEQ ID NO: 27:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9100 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Haemophilus influenzae  
 ; US-08-743-637B-27

Query Match 5.94; Score 117.2; DB 2; Length 9100;  
 Best Local Similarity 50.3%; Pred. No. 2e-25;  
 Matches 349; Conservative 0; Mismatches 333; Indels 12; Gaps 2;

OY 22 TCAACTATGCTGACTGGTCTTGAACGCCGCTCAATGCCCAAGCTAATGATATTC 81  
 |||||  
 DB 3498 TAAATTAATTCGGAAGGCGAGACGCCGATTCAGTGAATTAAGCCGATGTGC 3439  
 OY 82 CACAGATTGGTGAAGCAATCGTTTATGGAAGACCATGCGCTTCGACCAAGGG 141  
 |||||  
 DB 3438 ACAACGTTAATTAATGAGCAATTTTAGCAACGGAAGACAGTGTGTTTACGATCATACGG 3379

OY 142 GATTGATACATCGGTATCTGGAGCTTTCTTGGCAATCTGCAA--GCAATTCCT 198  
 |||||  
 DB 3378 ATTAAGACCTTATCGGATATGCCCCGTCATGTTTCCAGAGTAAATGGGGTGATC 3319  
 OY 199 CCAAGCTGATCAATCTCCACACAGTGTGATTAAGTTAGTACTTCTTCACTGCAC 258  
 |||||  
 DB 3318 ACAAGGCCAGTATGATTAATCAACAATTAAGCGGTAACTTTTCTTA-----AC 3268  
 OY 259 TTCCAGCAGACTATTTCTGTAAGCTCAGAACCTTGTAGCATTAAGTAAACA 318  
 |||||  
 DB 3267 CTCGAAAAAACCATTTATTCGTAAGCTCTGTAACCCGCTGTGGGAGTAAGTAAGAAA 3208  
 OY 319 AAAAGCAACCAAGCAAGAACTTGACCTACTATATTAATTAAGTCTACATGTAATGG 378  
 |||||  
 DB 3207 TACTCTCAACAACAAGAAATTAATAGCTTATTTAAACAAATCTTTTAAAGCTATCG 3148  
 OY 379 GAACATATGAATGACAGACGCTCAAAACACTATGTTAAAGACCTCAATTAATTAAG 438  
 |||||  
 DB 3147 TTCTTATGTTGGAGGGGAGCAAAACCTATTTGGTAAATCATTAATGAATTAATGAC 3088  
 OY 439 TTTACCTGAGTTAGCTTGTGCTGGAATGCTCAGGACCAACAATATGACCCCTA 498  
 |||||  
 DB 3087 CTATCGGAATGGGATTTATGCTGTTACTTAAGCACCTTCAACAATGAACCCGCT 3028  
 OY 499 TTCACATCAGAGACGACCCCAAGACCGCAAACTTGTCTTATCTGAATGAATAATCA 558  
 |||||  
 DB 3027 TTATCTTTAAACGTTGCAAGAACCGGCAATGTGCTAAACCCGATGTTAGATGA 2968  
 OY 559 AGGCTACATCTCTGCTGAACAGTATGAGAAGACACTCAATACCAATTAAGTGAAGACT 618  
 |||||  
 DB 2967 AAAATATCATCAGCAAGAAAGATATGATGCTGCAATGAAGACCGCATTTGGCGACTA 2908  
 OY 619 ACAAGTCTCAATCAGCAAGTAAATTAACCTGCTTACATGATTAATTCCTCAAGAGT 678  
 |||||  
 DB 2907 TCAGCGCCCAAAATTTGAATTTCCAGCCGATTTACTACTGAAATGGTGCACAGAAAT 2848  
 OY 679 CATCAATCAAGTTGAAGAAAGACGCTATTAAC 712  
 |||||  
 DB 2847 GGTGCGTGTGTTGGCGAAGAAATGCTTACACC 2814

RESULT 6

US-08-526-840B-27/c  
 ; Sequence 27, Application US/08526840B  
 ; Patent No. 6001564  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BERGERON, Michel G.  
 ; APPLICANT: OUELLETTE, Marc  
 ; APPLICANT: ROY, Paul H.  
 ; TITLE OF INVENTION: SPECIFIC AND UNIVERSAL PROBES AND  
 ; TITLE OF INVENTION: AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND IDENTIFY  
 ; TITLE OF INVENTION: COMMON BACTERIAL PATHOGENS AND ANTIBIOTIC RESISTANCE GENES  
 ; TITLE OF INVENTION: FROM CLINICAL SPECIMENS FOR ROUTINE DIAGNOSIS IN ...  
 ; NUMBER OF SEQUENCES: 177  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: O'ARLES & BRADY  
 ; STREET: 411 East Wisconsin Avenue  
 ; CITY: Milwaukee  
 ; STATE: Wisconsin  
 ; COUNTRY: USA  
 ; ZIP: 53202-4497  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/526,840B  
 ; FILING DATE: 11-SEP-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/304,732  
 ; FILING DATE: 12-SEP-1994





Db 200 GCTTCTAGAGACATACAGCGGTTGACCCGGTGGGATCTTCCTGACGACAGCGCTGGCGC 259  
QY 183 TGCAGAGCAATTCCTCCAGAGTGATCACTCTACCCCAACAGTTGATTAAGTTGACTT 242  
Db 260 TGTTCCTCCGCTACGCGCTCAGACAGGGCACTACCTATTACCCAGAGCTGGCGAGAACT 319  
QY 243 ACTTTCACTTCGACTTCGACGACGACTATTTCTGTAAGGCTCAGAGACCTTGGTTAG 302  
Db 320 TCTTCCTCAGTCC-----AGAACGCAAGCTGATGCGTAAGATTAAGGAAGCTTCCCTCG 373  
QY 303 CGATTCACTAGAACAAAAGCAACCAAGCAAGAAATCTTCACTACTATTAATAAGG 362  
Db 374 CGATTTCGATTGAACACTGCTGACGAAAGACGATCTCGACTTATCTGGAACAGA 433  
QY 363 TCTCATGCTTAATGGAAGACTGATGAATGACAGACAGCTCAACAACTACTATGTAAG 422  
Db 434 TTTACCTGGTTACCGGCTCATGTGTGCTGCTGGCGCAAGTCTATTTCCGAAAAA 493  
QY 423 ACCTCAATTAATTTAAGTTTACCTGATTAAGCTTGTGCTGGAATGCTCAGGACCA 482  
Db 494 CGGTGACCAACTGACGCTGACGAAGATGGCGGTGATAGCGGGCTGGCGAAAGCGCTT 553  
QY 483 ACCAATATGACCCCTATTACATCCAGAGACGCCCAAGCCGCAACTTGTCTTAT 542  
Db 554 CCACCTTCAACCCCTCTACTGATGATGCTGCCGCTGCGCGCTAAGCTCTGCTGT 613  
QY 543 CTGAATGAAGAAATCAAGGCTACATCTGCTGAACAGATGAGAAAGCACTCAATAAC 602  
Db 614 CGCGGATGCTGGATGAAGGATATATACCCCAACAGTTGATCAGACGACGACTGAGG 673  
QY 603 CAATTACTGAT 613  
Db 674 CGATTACGCT 684

RESULT 8  
US-08-771-716-1  
; Sequence 1, Application US/08771716  
; Patent No. 5922540  
; GENERAL INFORMATION:  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Hoskins, Joann  
; APPLICANT: Jaskunas, S.Richard  
; APPLICANT: Skatrud, Paul L.  
; TITLE OF INVENTION: Monofunctional Glycosyltransferase  
; TITLE OF INVENTION: Gene of Staphylococcus Aureus  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/771,716  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11067  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 807 base pairs  
; type: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..807  
US-08-771-716-1

Query Match 3.18; Score 62.2; DB 2; Length 807;  
Best Local Similarly 49.9%; Pred. No. 3,4e-09;  
Matches 191; Conservative 0; Mismatches 183; Indels 9; Gaps 1;

QY 67 AGCTAATGATATTCCTCCACAGATGTTGGTTAAGGCAATGCTTCTATGCAAGCAATTCCTT 126  
Db 252 AGCTGATTAACATGCGAGAGATGTTAAAGTGCTTTATTTCAATGGAAGTGAACGATT 311  
QY 127 CTTCGACCACAGGGGAGATTGATACCATCCGTATCTGGAGCTTCTTGGCAATTCGCA 186  
Db 312 CTACATATCATATGATTCGATTGTAAGATGACAACTAGACCTTATTTCAACGATTAG 371  
QY 187 AAGCAATTCCTCCAGAGTGATCACTTCACCAACAGTTGATTAAGTTGACTTACTT 246  
Db 372 CGACAGAGATGTCAGAGGTGTAGTACATTAACACAAAGTTTCAAAATTAATTTT 431  
QY 247 TTCACTTCGACTTCGACACGACGATTTTCGTAAGGCTGAGAGCGTTGATAGGAT 306  
Db 432 TGATTAAT-----GATGCTTATTACTGAAAGATTAAGAAATTAATTTGATGCTCA 482  
QY 307 TCAGTTAGAACAAAGCAACCAAGCAAGAAATCTTGACCTACTATTAATAAGGCTTA 366  
Db 483 TCGAGTTGAAAAACAATATATATAGAACGAATTTTAAGCTTTATTAATAATATTTTA 542  
QY 367 CATGCTAAGGAGACTATGGAATGACAGACGAGCTCAAACTACTATGTAAGACT 426  
Db 543 CTTGGGATTAATCAATATACGCTTGAGGGGCGAGCAAAACATTACTTGGAAACACGT 602  
QY 427 CAATTAATTAAGTTTACCTCAGT 449  
Db 603 GAATTAATATGTCACACATGT 625

RESULT 9  
US-08-771-716-3  
; Sequence 3, Application US/08771716  
; Patent No. 5922540  
; GENERAL INFORMATION:  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Hoskins, Joann  
; APPLICANT: Jaskunas, S.Richard  
; APPLICANT: Skatrud, Paul L.  
; TITLE OF INVENTION: Monofunctional Glycosyltransferase  
; TITLE OF INVENTION: Gene of Staphylococcus Aureus  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/771,716  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:

```

? NAME: Webster, Thomas D.
? REGISTRATION NUMBER: 39,872
? REFERENCE/DOCKET NUMBER: X-11067
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 317-276-3334
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 807 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: mRNA
? HYPOTHEetical: NO
? ANTI-SENSE: NO
? OS-08-771-716-3

```

Query Match	3.1%	Score 62.2	DB 2	Length 807
Best Local Similarity	33.9%	Pred. No. 3.4e-09		
Matches 130, Conservative	61	Mismatches 183	Indels 9	Gaps 1

OY	67	AGCTAAATGATTTTCCACAGATTGGTTAAGGCATGTTCATATCGAACATCATGGCTT	126
Db	252	AGCGTAACAACAGGCCAAGATATGUAAGAAGGCCUUAUUUCAUUGAGMAATAGACGAU	311
OY	127	CTTGACACACAGGGGATGTATACATCGTATCTGGAGCTTCTTCTGGCAATCTGCA	186
Db	312	CTACAUCAUCAUAGGAUUGCUUUUGAAGGUAACAUAAGAGCUUUUUUCAAACGAUAG	371
OY	187	AAGCAATTCCTCCAAGGTGGATCACTTCACCCACAAGACTTGATTAAGTTGACTTACTT	246
Db	372	CGACACAGAUUGUGCAGAGUGGUAGUACCAUACAACAAGAUUGUCAAAAUUAAUUUA	431
OY	247	TTCACATTCGCATTCGACCAGACGATATTTCTCGTAAGCGTCAGAGAACTGGTTACGAT	306
Db	432	UGAUAU-----GATCGUUCUADUUDACUAGAAAAGUAAAAGAUUUUUUAGUCUCA	482
OY	307	TCAGTTAGAACAAAAAGCAACCAAGCAAGAATCTTACCTACTATATATAATAAAGCTTA	366
Db	483	UCGAGUUGAAAAACAUAUAUUAAGAACGAUUUUUUAAGCUUUUUUUAAUUAAUUUA	542
OY	367	CATGTCTATGGAACATATGGAATGCAGACAGCAGCTCAAAACTACTATGTAAAGACT	426
Db	543	CUUUGGGAUUAUCAUAUAUACGCUUGAGGCGCACAAACCAUADACUUGGAAACAACCG	602
OY	427	CAATATTTAAGTTTACTCAGT	449
Db	603	GAAUAAAAAUAGUACAACAUAUG	625

```

10 RESULT 10
11 US-09-057-720A-1
12 : Sequence 1, Application US/09057720A
13 : Patent No. 6143868
14 :
15 : GENERAL INFORMATION:
16 :
17 : APPLICANT: Peery, Robert B.
18 :
19 : APPLICANT: Hoskins, Joann
20 :
21 : APPLICANT: Jaskunas, S.Richard
22 :
23 : APPLICANT: Skatrud, Paul L.
24 :
25 : TITLE OF INVENTION: Monofunctional Glycosyltransferase Gene
26 :
27 : TITLE OF INVENTION: of Staphylococcus Aureus
28 :
29 : NUMBER OF SEQUENCES: 3
30 :
31 : CORRESPONDENCE ADDRESS:
32 :
33 : ADDRESSEE: Eli Lilly and Company
34 :
35 : STREET: Lilly Corporate Center
36 :
37 : CITY: Indianapolis
38 :
39 : STATE: Indiana
40 :
41 : COUNTRY: U.S.
42 :
43 : ZIP: 46285
44 :
45 : COMPUTER READABLE FORM:
46 :
47 : MEDIUM TYPE: Floppy disk
48 :
49 : COMPUTER: IBM PC compatible
50 :
51 : OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

1      SOFTWARE: PatentIn Release #1.0, Version #1.30
2
3      CURRENT APPLICATION DATA:
4      APPLICATION NUMBER: US/09/057,720A
5
6      FILING DATE:
7
8      CLASSIFICATION: 435
9
10     ATTORNEY/AGENT INFORMATION:
11     NAME: Webster, Thomas D.
12
13     REGISTRATION NUMBER: 39,872
14
15     REFERENCE/DOCKET NUMBER: X-11067
16
17     TELECOMMUNICATION INFORMATION:
18
19     TELEPHONE: 317-276-3334
20
21     INFORMATION FOR SEQ ID NO: 1:
22
23     SEQUENCE CHARACTERISTICS:
24
25     LENGTH: 807 base pairs
26
27     TYPE: nucleic acid
28
29     STRANDEDNESS: single
30
31     TOPOLOGY: linear
32
33     MOLECULE TYPE: DNA (genomic)
34
35     HYPOTHETICAL: NO
36
37     ANTI-SENSE: NO
38
39     FEATURE:
40
41     NAME/KEY: CDS
42
43     LOCATION: 1..807
44
45     OS-09-057-720A-1

```

Query Match	3.1%;	Score 62.2;	DB 3;	Length 807;
Best Local Similarity	49.98;	Pred. No. 3.4e-09;		
Matches 191; Conservative	0;	Mismatches 183;	Indels 9;	Gaps 1

QY	67	AGCAATGATTTTCCCAACAGTTTTGGTTAAAGGCAATCGTTTCATGCACAACCATGGCTT	126
Db	252	AGCTGATTAACATGCCAGAGTAGTTGTTAAAGTGCTCTTATTTCATGTGAAGATGACGATT	311
QY	127	CTTGGACCACAGGGGAGTTGATTCACATCCGATTCCTGGAGACTTCTTGGCACAATCGCA	186
Db	312	CTACATCATCATGSGATTTCGATTTGAAAGGTACACAAGTACGTTTATTTCACACGATTAG	371
QY	187	AAGCAATTCCTCCCAAGSGTGANTCACTCACCCAACAGTTGATTAAGTTGACTTACTT	246
Db	372	CGACAGAGATGTGCAAGGTGGTAGTACCATTCATACACACAGTTGTCAAATAATTATTTTAA	431
QY	247	TTCACACTTCGACTTCGAGCCAGACGACTATTTTTCGTAAGGCTCAGSAGCTTGTTAGCGAT	306
Db	432	TGATTAAT-----CATCGTTCATTATTACTAGAAAAGTAATTAATTGTACTCA	482
QY	307	TCAGATTGAACAAAGAACAACGACGAAGAATCTTACCTACTATATAAATTAAGGCTTA	366
Db	483	TCGAGTTGAAAAAACATATATATATAAGAACGAAATTTTAAGCTTTTATTAATTAATTTTA	542
QY	367	CATGTCTAATGGGAATATGGAATGACAGACGACGCTCAAAACTACTTAATGTTAAGACCT	426
Db	543	CTTTGGGGATTAATCAATATATAGCGCTTGAGGGCGGACAAACCATTACTTTGGAAACAACCG	602
QY	427	CAATAATTTAAGTTACTCAGT	449
Db	603	GAAATAAAAATGTACAACTGT	625

RESULT 11  
 US-09-057-720A-3  
 Sequence 3, Application US/09057720A  
 Patent No. 6143868  
 GENERAL INFORMATION:  
 APPLICANT: Peery, Robert B.  
 APPLICANT: Hoskins, Joann  
 APPLICANT: Jaskunas, S. Richard  
 APPLICANT: Skarud, Paul L.  
 TITLE OF INVENTION: Monofunctional Glycosyltransferase Gene  
 TITLE OF INVENTION: of Staphylococcus Aureus  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Eli Lilly and Company

```

STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,720A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11067
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 807 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-057-720A-3

```

```

Query Match
Best Local Similarity 3.1%; Score 62.2; DB 3; Length 807;
Matches 130; Conservative 61; Mismatches 183; Indels 9; Gaps 1;

QY 67 AGCTAATGATTTCCACAGATTGGTTAAGCAATCGTTTCTATCGAAGACATCGCTT 126
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 252 AGCGAUAACAUGCCAGAGUAGUUAAGGUGCCUUUUAUUGCAUGGAAGUAGACG 311
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 CTGACACAGAGGGGATTTACATCCGTATCCCTGGAGCTTTCTGGCCATCTGCA 186
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 312 CUACAUAUCAUGAGUUGCAUUGAAAGUACAACUAGACUUUAUUGCAACAGUAG 371
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 AAGCAATTCCTCCAGAGTGATCTCCACCAACAGTTGATTAAGTTAGTTACTT 246
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 372 CGACAGAGAUUGCAGUGUAGUACAUUACAACAAGUUGUCAAUAUUUUUA 431
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 247 TTCACCTTCGACTCCGACAGACTATTCTCGTAGGCTCAGAGCTTGTTAGCGAT 306
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 432 UGAUAU-----GADCCUDCAUUDUAGAAAGUAAAGUAUUAUUGUAGCUCA 482
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 307 TCAGTTGAACAAAGCAACCAAGCAAGAAATCTTGACCTATATATAATAGGTCTA 366
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 483 UCGAGUUGAUAUAUAUAUAAGAAAGAAUUAAGCUUUUAUUAUUAUUAUUA 542
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 367 CATCTGAATGGAAGTGAATGACAGACAGCTCAAAACTACTATGTTAAAGACT 426
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 543 CUUUGGGAUUAUCAUUAUAGCUUGAGGGGCGACAAACAUUAUUUGAACAACCG 602
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 427 CAATAATTAATTACCTCAGT 449
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 603 GAUUAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 625
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 12
US-08-731-716-1
Sequence 1, Application US/08731716
Patent No. 5789202
GENERAL INFORMATION:
APPLICANT: Hoskins, Joan
APPLICANT: Jaskunas, S. Richard
APPLICANT: Rocket, Pamela K.

```

```

APPLICANT: Zhao, Genshi
APPLICANT: Rostock, Paul R. Jr.
APPLICANT: No. 5789202, Franklin H.
TITLE OF INVENTION: Penicillin Binding Protein From
TITLE OF INVENTION: Streptococcus Pneumoniae
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/731,716
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-10,887
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2193 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2193
US-08-731-716-1

```

```

Query Match
Best Local Similarity 2.8%; Score 56.4; DB 1; Length 2193;
Matches 381; Conservative 0; Mismatches 446; Indels 27; Gaps 3;

QY 5 ATTCAGCAATTAATCAATCTATTCCTGCTTGGGTTGTAACCGCGCTCAATGCC 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 283 ATTTTGAACCGTGAAGAAAGAGGCTGCTCTGTCTGTCAAAGGAACTATGTT 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 65 CAAGCTAATGATTTCCACAGATTGGTTAAGCAATCGTTTCTATGAGACCATGCG 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 343 GAGCTGACTGCATCAATGAATTAAGTCTGTTATTTGGACAGAAAGACGTTCT 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 125 TTCTTCGACACAGGGGATTTGATACATCCGTAATCTGGAGGCTTTCTTGGAATCTG 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 403 TTCTATAAATGACGGGATTAACATAGCGCTTC-----TTCTTGATATGTC 453
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 185 CAAGCAATTCCTCCCAAGTGAATCAACTCTCACCCCAACAGTTGATTAAGTTAC 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 454 ACTGCTGAGCGTTGAGGTGGCTCTACATTAACCAACAGCTGCTGAATAAAGCGCTAT 513
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 245 TTTCACCTTGACCTCCGACCAAGATTTCTGTAAGGCTCAGAAAGCTTGTTAGCG 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 514 TTAT-----CGAGGATCAAACTGTTGAGAGAAAGGAAAGATTTTCTTGCC 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 305 ATTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTTGACCTACTATTAATTAAGTGC 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 565 TTAGATTTAAGCAAAAATATAGTAAGAGCAATTTCTAACCATGTAACCAAGCT 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 365 TACATGCTTAATGGAACTATGATGACAGACAGCTCAAAACTACTATGTTAAAGAC 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```
RESULT 14
US-09-060-756-206
; Sequence 206, Application US/09060756
; Patent No. 6183957
; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
; APPLICANT: Buchrieser-Brosch, Roland
; APPLICANT: Gordon, Stephen
; APPLICANT: Billault, Alain
; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
; TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
; FILE REFERENCE: 3495-0169
; CURRENT APPLICATION NUMBER: US/09/060,756
; CURRENT FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 206
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (various positions within the sequence)
; OTHER INFORMATION: applicants are uncertain of bases designated as "n"
US-09-060-756-206
```

```
Query Match 2.88; Score 55.2; DB 4; Length 428;
Best Local Similarity 50.88; Pred. No. 3.3e-07;
Matches 132; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 281 AAGCTCAGAGCTGGTGTACGCTTACGTTAGAACAAAAGAACCAACGACGAAGAATC 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 aagcttcgagatccgagatgacacacgctgacaaagacccacaaatctgaatc 66

QY 341 TTGACCTACTATTAATTAATAGCTTACATGTCTAATGGGAATATGGAATCGACACACA 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 ctgacccgatacttgacactgctcgtctgcaataactcgttcgagcgagcagcg 126

QY 401 GCTAAACTACTATGCTAAGAAAGACCTCAATATTTAAGTTTAACTTACCTGAGCTTGCTG 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 gcgcaaacgctacttcgacatcaacgctcgcgacctgaatgacgcaagcgcgctgctg 186

QY 461 GCTGAATGCTTCAGGACCAACCAATATGACCCCTATTACATCCAGAGCAAGCCCAA 520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 gcgagcatggtgcaatcgaccagcagctcaacccgtacacaccccgacgagcgctg 246

QY 521 GACCGCGAAACTGTGCTT 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 gcccgcggaacgctgctct 266
```

```
RESULT 15
US-09-508-542-17
; Sequence 17, Application US/09508542
; Patent No. 6339174
; GENERAL INFORMATION:
; APPLICANT: STRAUS, ANDREAS
; APPLICANT: THUMM, GUNTHER
; APPLICANT: POHLNER, JOHANNES
; APPLICANT: GÖTZ, FRIEDRICH
; TITLE OF INVENTION: METHOD FOR IDENTIFYING A NUCLEIC ACID
; FILE REFERENCE: 10496/P65266USO
; CURRENT APPLICATION NUMBER: US/09/508,542
; CURRENT FILING DATE: 2000-05-16
; PRIOR APPLICATION NUMBER: PCT/EP98/06136
; PRIOR FILING DATE: 1998-09-26
; PRIOR APPLICATION NUMBER: 97 116 841.4
; PRIOR FILING DATE: 1997-09-27
; PRIOR APPLICATION NUMBER: 97 118 755.4
```

```
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Staphylococcus carnosus
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(1474)
; OTHER INFORMATION: "n" represents a, t, c, g, unknown or other
US-09-508-542-17
```

```
Query Match 2.78; Score 53.8; DB 4; Length 1474;
Best Local Similarity 50.48; Pred. No. 1.7e-06;
Matches 183; Conservative 0; Mismatches 174; Indels 6; Gaps 2;

QY 1285 AAAGATGGCTGCTGCTTACGCTGCTTGGCAATGTTGGAACCTTACTATTAACCAATGTA 1344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 766 acaattagcatcagcatctgctgcatcgcgaacgltgctactatacaacgagcatlc 825

QY 1345 TATCCATTAAGTCTGCTTTAGTATGATGGAGTGAAGAAAGTTCTCTAATGTGGAACTCG 1404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 826 aattcaaaaagtagtactcgtggaacnltgaaacaaatcgaatcagatcactaagcataa 885

QY 1405 TGCCATGAAGGAAGAACAGACAGCCTATATGATGACGACATGATGA--ACAGTCTTGAC 1461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 886 agcgatgagtgatatacactcgtcatcatcagtgctgagatgctaaaggtactaataac 945

QY 1462 TTATGGAACGTGACGAATGCTTATCTTGTGCTCCCTCAGCCTGGTAAACAGAAC 1521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 946 ataggtctcgtcatatgccaatgltgctgagatlaatatltgtctaaagacaggtac 1005

QY 1522 CTCATACATATACAGACGAGGAATTTGAAACCATCATCAAGACTCTCATTTGTAGACAC 1581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1006 tgttacttac--gtgtcgaaactatctcaaatataatacttaccgtataatgagcga 1062

QY 1582 TGATGAACCTATTGCTGCTTATGCGCTAATATTTCAATGGCTGTATGACAGGCTATTTC 1641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1063 agcagtggtgattaaagcgttataccctcaatacactatgcaatgltgagatggtctcag 1122

QY 1642 TAA 1644
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1123 taa 1125
```

Search completed: June 13, 2002, 13:45:41  
Job time: 4168 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 12:37:03 ; Search time 3441.19 Seconds  
(without alignments)  
12567.793 Million cell updates/sec

Title: US-08-961-083-1

Perfect score: 1999

Sequence: 1 TAAATCTAGACATTAATAA.....ATCCTCAACGACGACAACCA 1999

Scoring table: IDENTITY\_NUC  
Gap 10.0 , Gape 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US080\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq.\*  
12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq.\*  
13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq.\*  
14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq.\*  
15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq.\*  
16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq.\*  
17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq.\*  
18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq.\*  
19: /cgn2\_6/ptodata/2/pna/US095\_COMB.seq.\*  
20: /cgn2\_6/ptodata/2/pna/US095b\_COMB.seq.\*  
21: /cgn2\_6/ptodata/2/pna/US095c\_COMB.seq.\*  
22: /cgn2\_6/ptodata/2/pna/US095d\_COMB.seq.\*  
23: /cgn2\_6/ptodata/2/pna/US096a\_COMB.seq.\*  
24: /cgn2\_6/ptodata/2/pna/US096b\_COMB.seq.\*  
25: /cgn2\_6/ptodata/2/pna/US096c\_COMB.seq.\*  
26: /cgn2\_6/ptodata/2/pna/US096d\_COMB.seq.\*  
27: /cgn2\_6/ptodata/2/pna/US096e\_COMB.seq.\*  
28: /cgn2\_6/ptodata/2/pna/US097a\_COMB.seq.\*  
29: /cgn2\_6/ptodata/2/pna/US097b\_COMB.seq.\*  
30: /cgn2\_6/ptodata/2/pna/US097c\_COMB.seq.\*  
31: /cgn2\_6/ptodata/2/pna/US098a\_COMB.seq.\*  
32: /cgn2\_6/ptodata/2/pna/US098b\_COMB.seq.\*  
33: /cgn2\_6/ptodata/2/pna/US098c\_COMB.seq.\*  
34: /cgn2\_6/ptodata/2/pna/US099a\_COMB.seq.\*  
35: /cgn2\_6/ptodata/2/pna/US099b\_COMB.seq.\*  
36: /cgn2\_6/ptodata/2/pna/US099c\_COMB.seq.\*  
37: /cgn2\_6/ptodata/2/pna/US101\_COMB.seq.\*  
38: /cgn2\_6/ptodata/2/pna/US101\_COMB.seq.\*  
39: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq.\*  
40: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq.\*  
41: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq.\*  
42: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq.\*  
43: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq.\*

44: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq.\*  
45: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq.\*  
46: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq.\*  
47: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq.\*  
48: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq.\*  
49: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq.\*  
50: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq.\*  
51: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq.\*  
52: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq.\*  
53: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq.\*  
54: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq.\*  
55: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq.\*  
56: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq.\*  
57: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq.\*  
58: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq.\*  
59: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq.\*  
60: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq.\*  
61: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq.\*  
62: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq.\*  
63: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq.\*  
64: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq.\*  
65: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq.\*  
66: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq.\*  
67: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq.\*  
68: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq.\*  
69: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq.\*  
70: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq.\*  
71: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq.\*  
72: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq.\*  
73: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq.\*  
74: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq.\*  
75: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	20	US-09-536-784-1
2	1999	100.0	1999	30	US-09-765-271-1
3	1999	100.0	1999	30	US-09-765-272-1
4	1999	100.0	10383	41	US-60-029-960-81
5	1999	100.0	10711	13	US-08-961-527-145
6	1987.8	99.4	2160	22	US-09-583-110-1312
7	1987.8	99.4	2166	15	US-09-107-433-1102
8	1981.4	99.1	2160	31	PCR-US02-03987-9325
9	1981.4	99.1	2160	31	US-09-815-242-9325
10	1981.4	99.1	2160	37	US-10-072-851-9325
11	1981.4	99.1	10333	45	US-60-061-998-596
12	1631.2	81.6	9845	45	US-60-068-175-596
13	944.2	47.2	960	5	US-08-116-541-3
14	808.6	40.5	2172	13	PCR-US97-14436-139
15	808.6	40.5	2172	13	US-08-911-503-139
16	808.6	40.5	2172	13	US-08-911-503A-139
17	574.6	28.7	2337	1	PCR-US02-03987-6631
18	574.6	28.7	2337	31	US-09-815-242-6631
19	574.6	28.7	2337	37	US-10-072-851-6631
20	573	28.7	2370	15	US-09-134-000-1534
21	559	28.0	8395	14	US-09-070-927-217
22	559	28.0	8395	14	US-09-070-927A-217
23	553	27.7	2472	15	US-09-107-532-2013
24	553	27.7	2472	15	US-09-107-532A-2013
25	496.2	24.8	1386	37	US-10-091-007-91
26	410	20.5	2470	44	US-60-068-186-495
27	410	20.5	2474	44	US-60-050-444-495
28	352.2	17.6	3478	45	US-60-068-217-710
29	348.6	17.4	3279	44	US-09-634-238-44
30	348.6	17.4	7818	24	US-09-634-238-14
31	273	13.7	933	43	US-60-045-649-108

Sequence 1, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 81, Appl  
Sequence 145, Appl  
Sequence 1312, Ap  
Sequence 9325, Ap  
Sequence 9325, Ap  
Sequence 596, App  
Sequence 596, App  
Sequence 3, Appl  
Sequence 139, App  
Sequence 139, App  
Sequence 139, App  
Sequence 6631, Ap  
Sequence 6631, Ap  
Sequence 6631, Ap  
Sequence 1534, Ap  
Sequence 217, App  
Sequence 217, App  
Sequence 2013, Ap  
Sequence 2013, Ap  
Sequence 91, Appl  
Sequence 495, App  
Sequence 495, App  
Sequence 710, App  
Sequence 710, App  
Sequence 14, Appl  
Sequence 108, App

C	32	273	13.7	933	43	US-60-046-653-193	Sequence 183, App
C	33	258	12.9	323	41	US-60-029-960-934	Sequence 934, App
C	34	221.4	11.1	6394	26	US-09-663-779-1130	Sequence 1130, App
C	35	204	10.2	204	22	US-09-583-110-1313	Sequence 1313, App
C	36	164.2	8.2	3748	26	US-09-663-779-1038	Sequence 1038, App
C	37	163	8.2	2229	36	US-09-974-300-1671	Sequence 1671, App
C	38	123.8	6.2	1899	54	US-09-620-608-1543	Sequence 1543, App
C	39	123.8	6.2	1899	53	US-60-144-883-1543	Sequence 1543, App
C	40	123.8	6.2	2568	17	US-09-138-352-2090	Sequence 2090, App
C	41	122.8	6.1	1590	42	US-60-038-697-335	Sequence 335, App
C	42	122.8	6.1	1590	43	US-60-046-714-308	Sequence 308, App
C	43	120.6	6.0	2181	1	PCT-US02-03987-4727	Sequence 4727, App
C	44	120.6	6.0	2181	31	US-09-815-242-4727	Sequence 4727, App
C	45	120.6	6.0	2181	37	US-10-072-851-4727	Sequence 4721, App

## ALIGNMENTS

```

RESULT 1
US-09-536-784-1
; Sequence 1, Application US/09536784
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
; US-09-536-784-1

```

	Query Match	100.0%	Score 1999:	DB 20:	Length 1999:
	Best Local Similarity	100.0%	Pred. No. 0:		
	Matches 1999:	Conservative	0:	Mismatches	0:
				Indels	0:
				Gaps	0:
QY	1 TAAATCTACGACATATAAAATCACTCATTTGGTCTCTGAACGCCGGCTCA	60			
Db	1 TAAATCTACGACATATAAAATCACTCATTTGGTCTCTGAACGCCGGCTCA	60			
QY	61 TGCCCAACTAATGATATCCACAGATTGGTTAAGCAATCGTTTCAATGAGACCA	120			
Db	61 TGCCCAACTAATGATATCCACAGATTGGTTAAGCAATCGTTTCAATGAGACCA	120			

OY	121	TCGCTCTTCGCACACAGGGGGAGTGAATACATCCGCTATCTCGGAGCTTCCTTCGGAA	180
OY	121	TCGCTCTTCGCACACAGGGGGAGTGAATACATCCGCTATCTCGGAGCTTCCTTCGGAA	180
OY	181	TCGCAAGAAGCAATTCCTCCAGGTGATCACTCTCACCACAAGTTGATTAAGTTGAC	240
Db	181	TCGCAAGAAGCAATTCCTCCAGGTGATCACTCTCACCACAAGTTGATTAAGTTGAC	240
OY	241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTT	300
Db	241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTT	300
OY	301	AGCGATTACGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACTACTATATTAATA	360
Db	301	AGCGATTACGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACTACTATATTAATA	360
OY	361	GGCTACATGTCATATGGGAGACTATGGAAATGCAGACAGACTCAAACTACTATGTAA	420
Db	361	GGCTACATGTCATATGGGAGACTATGGAAATGCAGACAGACTCAAACTACTATGTAA	420
OY	421	AGACCTCAATATTAATTAGTGTACCTAGTTAGGCTTGCGGTGGAAATCCCTCAGGACC	480
Db	421	AGACCTCAATATTAATTAGTGTACCTAGTTAGGCTTGCGGTGGAAATCCCTCAGGACC	480
OY	481	AAACCAATATGACCCCTCATTCACATCCAGAGCAGCCCAAGACCAGCAACTTGCTTT	540
Db	481	AAACCAATATGACCCCTCATTCACATCCAGAGCAGCCCAAGACCAGCAACTTGCTTT	540
OY	541	ATCTGGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAATAC	600
Db	541	ATCTGGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAATAC	600
OY	601	ACCAATTAAGTAAATATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAATAC	660
Db	601	ACCAATTAAGTAAATATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAATAC	660
OY	661	TAAATTAAGTAAATATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAATAC	720
Db	661	TAAATTAAGTAAATATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAATAC	720
OY	721	AACGGGATGGATGTCACCAAAATGAGAACCAAGGCTCAAAAACCTCTGGGATAT	780
Db	721	AACGGGATGGATGTCACCAAAATGAGAACCAAGGCTCAAAAACCTCTGGGATAT	780
OY	781	TTACAATACAGACGAATAGCTTCCTATCCAGACGATGGAATTCGAAGTGCCTTACCAT	840
Db	781	TTACAATACAGACGAATAGCTTCCTATCCAGACGATGGAATTCGAAGTGCCTTACCAT	840
OY	841	TGTTGATGTTTCTAAGGTAAATCAATGCGCCAGTAGAGGACGCATCAAGTCAAGTAA	900
Db	841	TGTTGATGTTTCTAAGGTAAATCAATGCGCCAGTAGAGGACGCATCAAGTCAAGTAA	900
OY	901	TGTTTCCCTGGGAATTAACCAAGCAGTAGAAACAAACCCGACTGGGATCAACTATGAA	960
Db	901	TGTTTCCCTGGGAATTAACCAAGCAGTAGAAACAAACCCGACTGGGATCAACTATGAA	960
OY	961	ACCGATACAGACTATGCTCTCTGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT	1020
Db	961	ACCGATACAGACTATGCTCTCTGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT	1020
OY	1021	CGTTACAGATGAGCCCTATATACCTCTGGGACAAATATCTCTGTTTATTAAGTGGATAG	1080
Db	1021	CGTTACAGATGAGCCCTATATACCTCTGGGACAAATATCTCTGTTTATTAAGTGGATAG	1080
OY	1081	GGGCTACTTTGGCAACATACCTTTGCAATACGCCCTGCAACAACTCGCAAGAGTCCAGC	1140
Db	1081	GGGCTACTTTGGCAACATACCTTTGCAATACGCCCTGCAACAACTCGCAAGAGTCCAGC	1140
OY	1141	CGTGGAAATCTTAACAAGGTGGGACTCAACCGCGCAAGACTTTTCTTAATAGTGTAGG	1200
Db	1141	CGTGGAAATCTTAACAAGGTGGGACTCAACCGCGCAAGACTTTTCTTAATAGTGTAGG	1200



QY	1201	AATCGATACCACCAAGTATCTCAGTACTGCAATATGCCATTTCCAGTAACACAAACCGAATCAGA	1260
Db	1201	AATCGATACCACCAAGTATCTCAGTACTGCAATATGCCATTTCCAGTAACACAAACCGAATCAGA	1260
QY	1261	CAAAAAATATGAGAGCAAGTATAGTGAAGAAAGATGGCTGCTGCTTAACGCTGGCTTTGGCAAAATGG	1320
Db	1261	CAAAAAATATGAGAGCAAGTATAGTGAAGAAAGATGGCTGCTGCTTAACGCTGGCTTTGGCAAAATGG	1320
QY	1321	TGCAAACTTACTATTAACCAATGTATATCCATTAAGTCTGCTTTTAACTGATGGAGATGAAA	1380
Db	1321	TGCAAACTTACTATTAACCAATGTATATCCATTAAGTCTGCTTTTAACTGATGGAGATGAAA	1380
QY	1381	AGAGTTCTCTAATATGTCGGCAACTGCTGCCATGGAAGAAACGACAGCCTTATATGATGACCGA	1440
Db	1381	AGAGTTCTCTAATATGTCGGCAACTGCTGCCATGGAAGAAACGACAGCCTTATATGATGACCGA	1440
QY	1441	CATGATTAACCAACAGCTCTTTCAGTATATGAGCACTGGAGCAATATGCTTATCTTGGCTGGCTCC	1500
Db	1441	CATGATTAACCAACAGCTCTTTCAGTATATGAGCACTGGAGCAATATGCTTATCTTGGCTGGCTCC	1500
QY	1501	TCAAGCTGTGTAAACAGGAACCTCTTAACCTATACAGACGAGAAATTTGAAAAACCAATCAAA	1560
Db	1501	TCAAGCTGTGTAAACAGGAACCTCTTAACCTATACAGACGAGAAATTTGAAAAACCAATCAAA	1560
QY	1561	GACCTCTCAATTTGTATGACACCTGATGATGAATATTTGCTGGCTTATACGGCTAAATATTTCAAT	1620
Db	1561	GACCTCTCAATTTGTATGACACCTGATGATGAATATTTGCTGGCTTATACGGCTAAATATTTCAAT	1620
QY	1621	GGCGTATGAGACAGGCGTATTTCAACGCTGTGACACCACTGTGAGCGAATGGCGCTTACGGT	1680
Db	1621	GGCGTATGAGACAGGCGTATTTCAACGCTGTGACACCACTGTGAGCGAATGGCGCTTACGGT	1680
QY	1681	CGCTGCCAAAGTTTACCGCTCTATGATGATGACACTGCTGTGAAGGAAACCAATCCAGAA	1740
Db	1681	CGCTGCCAAAGTTTACCGCTCTATGATGATGACACTGCTGTGAAGGAAACCAATCCAGAA	1740
QY	1741	TTGGAATATATCCAGAGGGGCTCTACAGAAATGGAGATTTGTAATTTAAAAATGGTGCTCG	1800
Db	1741	TTGGAATATATCCAGAGGGGCTCTACAGAAATGGAGATTTGTAATTTAAAAATGGTGCTCG	1800
QY	1801	TTTACGTGGAACCTCACTGCTGTCCACAAACACCCCATCACTGAATGTAAGTTCACACTATC	1860
Db	1801	TTTACGTGGAACCTCACTGCTGTCCACAAACACCCCATCACTGAATGTAAGTTCACACTATC	1860
QY	1861	ATCAGATAGTTCACATTCACAGTCTAGTCAACCACTCCAGACCAATTAATAGTAGAC	1920
Db	1861	ATCAGATAGTTCACATTCACAGTCTAGTCAACCACTCCAGACCAATTAATAGTAGAC	1920
QY	1921	TACCAATCTCAACATATATACGACACATCAATCAAAATCAACCCCTGATCAACAAAAATTCAGAA	1980
Db	1921	TACCAATCTCAACATATATACGACACATCAATCAAAATCAACCCCTGATCAACAAAAATTCAGAA	1980
QY	1981	TTCTCAACGACGACCAACCA	1999
Db	1981	TTCTCAACGACGACCAACCA	1999
RESULT 2			
US-09-765-271-1			
Sequence 1, Application US/09765271			
GENERAL INFORMATION:			
APPLICANT: Choi et. al.			
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines			
NUMBER OF SEQUENCES: 452			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Human Genome Sciences, Inc.			
STREET: 9410 Key West Avenue			
CITY: Rockville			
STATE: Maryland			
COUNTRY: USA			
ZIP: 20850			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage			

```

1 COMPUTER: HP Vectra 486/33
2 OPERATING SYSTEM: MSDOS version 6.2
3 SOFTWARE: ASCII Text
4
5 CURRENT APPLICATION DATA:
6 APPLICATION NUMBER: US/09/765,271
7 FILING DATE: 22-Jan-2001
8
9 CLASSIFICATION: <Unknown>
10
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: 09/536,784
13 FILING DATE: <Unknown>
14 APPLICATION NUMBER: 08/961,083
15 FILING DATE: OCT-30-1997
16
17 ATTORNEY/AGENT INFORMATION:
18 NAME: Michelle S. Marks
19 REGISTRATION NUMBER: 41,971
20 REFERENCE/DOCKET NUMBER: PB340P3
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: (301) 309-8504
23 TELEFAX: (301) 309-8512
24
25 INFORMATION FOR SEQ ID NO: 1:
26
27 SEQUENCE CHARACTERISTICS:
28 LENGTH: 1999 base pairs
29 TYPE: nucleic acid
30 STRANDEDNESS: double
31 TOPOLOGY: linear
32
33 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
34 US-09-765-271-1

```

Query Match	100.0%	Score 1999:	DB 30:	Length 1999:
Best Local Similarity	100.0%	Pred. No. 0:		
Matches 1999: Conservative	0:	Mismatches	0:	Indels
				Gaps
QY 1 TAAATCTACGACAAATATAAAATCAACTATGCTACTGTGGGTCTCGAAGCGCGGTCAA	60			
Db 1 TAAATCTACGACAAATATAAAATCAACTATGCTACTGTGGGTCTCGAAGCGCGGTCAA	60			
QY 61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCATCGTTTCTATCGAAGACCA	120			
Db 61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCATCGTTTCTATCGAAGACCA	120			
QY 121 TCGCTTTCTCGACACAGGGGGATGTATACATCCGTAATCCGGAAGCTTTCTTGCGCAA	180			
Db 121 TCGCTTTCTCGACACAGGGGGATGTATACATCCGTAATCCGTAATCCGGAAGCTTTCTTGCGCAA	180			
QY 181 TCTGCAAGCAATTCCTCCACAGGTGGATCAACTCTCAACCCCAAGTGTATTAAGTTGAC	240			
Db 181 TCTGCAAGCAATTCCTCCACAGGTGGATCAACTCTCAACCCCAAGTGTATTAAGTTGAC	240			
QY 241 TTACTTTTCAACTTCGACTTCGACACACTTTTCTGTAAAGCTTCAGGAAGCTTGGT	300			
Db 241 TTACTTTTCAACTTCGACTTCGACACACTTTTCTGTAAAGCTTCAGGAAGCTTGGT	300			
QY 301 AGCGATTCAGTTAGAACAAAAAGCAACGACGAAGAAATCTTGACCTACTATATAATTA	360			
Db 301 AGCGATTCAGTTAGAACAAAAAGCAACGACGAAGAAATCTTGACCTACTATATAATTA	360			
QY 361 GGTCTACATGCTCTAATGGAAGTATGGAATGCAGACGACGCTCAAACTACTATGTTAA	420			
Db 361 GGTCTACATGCTCTAATGGAAGTATGGAATGCAGACGACGCTCAAACTACTATGTTAA	420			
QY 421 AGACCTCAATATTAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAA	480			
Db 421 AGACCTCAATATTAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAA	480			
QY 481 AAACCAATATGACCCCTATTTCATCTCAGAAAGACGCCCAAGACCTTGCTTT	540			
Db 481 AAACCAATATGACCCCTATTTCATCTCAGAAAGACGCCCAAGACCTTGCTTT	540			
QY 541 ATCTGAATGAAAAATCAAGGCTAAATCTCTGCTGAACAGTATGAAAAAGCATTAATAC	600			
Db 541 ATCTGAATGAAAAATCAAGGCTAAATCTCTGCTGAACAGTATGAAAAAGCATTAATAC	600			

OY	601	ACCAATTACGATGAGACTTACAAAGTCCAAATACGAACTAATTTCACCTGGCTACATATGA	660
Db	601	ACCAATTACTGATGAGACTACAAAGTCTCAAATACGAACTAATTTCACCTGGCTACATATGA	660
OY	661	TAATTACCTCAAGCAAGTCTCATCATCACTGAGTGAAGAAGCAAGGCTATTAACTTACTCAC	720
Db	661	TAATTACCTCAAGCAAGTCTCATCATCACTGAGTGAAGAAGCAAGGCTATTAACTTACTCAC	720
OY	721	AACGTGGATGGATGTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
Db	721	AACGTGGATGGATGTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
OY	781	TTTCAATACAGACGAATAAGCTTGCGTATCCAGACGATGTAATTGCAAGTGCCTTACCAT	840
Db	781	TTTCAATACAGACGAATAAGCTTGCGTATCCAGACGATGTAATTGCAAGTGCCTTACCAT	840
OY	841	TGTTGATGTTCTTAAACGCTAAAGTCAATTGGCCAGCTAGAGACGCCATCACTAGTCAAGTAA	900
Db	841	TGTTGATGTTCTTAAACGCTAAAGTCAATTGGCCAGCTAGAGACGCCATCACTAGTCAAGTAA	900
OY	901	TGTTTCTTGGGAATTACCAAGCACTAGAAACAACCCGCACTGGGATCACTATGAA	960
Db	901	TGTTTCTTGGGAATTAAACCAACAGTAGAAACAACCCGCACTGGGATCAACTATGAA	960
OY	961	ACCGATCAAGACATATGCTCTGCGTTGGAGTACGGGTCTACGATTAACACTGACTAT	1020
Db	961	ACCGATCAAGACATATGCTCTGCGTTGGAGTACGGGTCTACGATTAACACTGACTAT	1020
OY	1021	CGTTACAGATAGAGCCCTATATCTAACCTCGGAGCAATACTCCTGTTTAACTGGGATAG	1080
Db	1021	CGTTACAGATAGAGCCCTATATCTAACCTCGGAGCAATACTCCTGTTTAACTGGGATAG	1080
OY	1081	GGGCTACTTTGGCAACATCACTTGCAATACGCCCTGCACAACATGCGAAACGTCCAGC	1140
Db	1081	GGGCTACTTTGGCAACATCACTTGCAATACGCCCTGCACAACATGCGAAACGTCCAGC	1140
OY	1141	CGTGGAAACTCTAAACAAGSTCGGACTCAACCGCGCAGAGACTTTCCTAATGGTCTAGG	1200
Db	1141	CGTGGAAACTCTAAACAAGSTCGGACTCAACCGCGCAAGCTTTCCTAATGGTCTAGG	1200
OY	1201	AATGACATACCCAAAGTWTACTACTACACAATGCGATTTCACTAACACACACCGAATCGA	1260
Db	1201	AATGACATACCCAAAGTWTACTACTACACAATGCGATTTCACTAACACACACCGAATCGA	1260
OY	1261	CAAAAAATATGGAGCAAGTAGTGAAGAAAGATGGCTGCTGCTTACGCTGGCTTGCAAATGG	1320
Db	1261	CAAAAAATATGGAGCAAGTAGTGAAGAAAGATGGCTGCTGCTTACGCTGGCTTGCAAATGG	1320
OY	1321	TGGAACTTACTATTAACCAATGTATATCCATAAAGTCGCTTATAGTATGGAGTGAAGA	1380
Db	1321	TGGAACTTACTATTAACCAATGTATATCCATAAAGTCGCTTATAGTATGGAGTGAAGA	1380
OY	1381	AGAATTCCTATATGTGCGAAGCTGCTGCCAATGGAAGAAAGCAGCGCTATATGATGACCGA	1440
Db	1381	AGAATTCCTATATGTGCGAAGCTGCTGCCAATGGAAGAAAGCAGCGCTATATGATGACCGA	1440
OY	1441	CATGATGAAAAACAGTCTTGAATTATGGAACGTGACGAAATGGCTATCTTGTGGCTGCC	1500
Db	1441	CATGATGAAAAACAGTCTTGAATTATGGAACGTGACGAAATGGCTATCTTGTGGCTGCC	1500
OY	1501	TCAGAGCTGCTAATGCGAAGCTCTTAACCTATACAGACGAGGAATTTGAAACACATACAA	1560
Db	1501	TCAGAGCTGCTAATGCGAAGCTCTTAACCTATACAGACGAGGAATTTGAAACACATACAA	1560
OY	1561	GACCTGCAATTTGTACACCTGATGAACATTTGTGGCTATACGCGTAATATTTCAAT	1620
Db	1561	GACCTGCAATTTGTACACCTGATGAACATTTGTGGCTATACGCGTAATATTTCAAT	1620
OY	1621	GGCTGTATGACAGAGCTATTTCAACCGTGTGACACACACTTGAAGCAATGGCTTACAGT	1680
Db	1621	GGCTGTATGACAGAGCTATTTCAACCGTGTGACACACACTTGAAGCAATGGCTTACAGT	1680
OY	1681	CGCTGCCAAAGTTTACCGCTATGATGACTACTCTGTCTGAAGAGCAATCCAGAAGA	1740

Db	1681	CGCTGCCAAGTTTACGGCTCTATGATGACCTACGTGCTGAAGGAAAGCAATCCAGAAAG	17408
Qy	1741	TTGGATATATACACAGAGGGGCTTACACAAATGAGAAATTCGATTTAAAAATGGTGCTCG	18000
Db	1741	TTGGAAATATACCAAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAAAATGGTGCTCG	18000
Qy	1801	TTTACAGGGAAACACCTCTGCTCCACAACACCCCATCACTACCTGAAAGTTCAAAGCTCATC	18666
Db	1801	TTTACAGGGGAACCTCTGCTCTCCACACACCCCATCACTAGAAAGTTCAAAGCTCATC	18666
Qy	1861	ATCAGATAGTTCAACTTCACAGCTCTAGCTTCACCACTCCAGACACAAATATATAGTAGAC	19200
Db	1861	ATCAGATAGTTCAACTTCACAGCTCTAGCTTCACCACTCCAGACAAATATATAGTAGAC	19200
Qy	1921	TACCAATCTTACCATATATACGCAACACATCTAAATACAACCCCTGATCAACAAAATACAGA	19800
Db	1921	TACCAATCTTACCATATATACGCAACACATCTAAATACAACCCCTGATCAACAAAATACAGA	19800
Qy	1981	TCCTCAACGACACACACCA 1999	
Db	1981	TCCTCAACGACACACCA 1999	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

D	b	1	TAAATCTACGACAACTATAAATATCACTCATCTGCTGACTTGGGTTCTGAAAGCCCGGCTCAA	60
O	y	61	TGCCCAAGCTATGATATATTCCTCCACAGATTTTGTTAAAGCAATCGTTTCTATCGAAGACCA	120
D	b	61	TGCCCAAGCTATGATATATTCCTCCACAGATTTTGTTAAAGCAATCGTTTCTATCGAAGACCA	120
O	y	121	TGCGTCTCTGACACACAGGGGGATGATACCATTCCGATCTCTGGAGCTTTCTTGCGCAA	180
D	b	121	TGCGTCTCTGACACACAGGGGGATGATACCATTCCGATCTCTGGAGCTTTCTTGCGCAA	180
O	y	181	TCTGCAAAAGCAATTCCTCCAGGTGGATACATCTCACCCAAACAGTGTATTAAGTGCAC	240
D	b	181	TCTGCAAAAGCAATTCCTCCAGGTGGATACATCTCACCCAAACAGTGTATTAAGTGCAC	240
O	y	241	TTACTTTTCAACTTCGACTTCGCCACAGACTATTCTCGTAAGGCTCAGAAAGCTTGGTT	300
D	b	241	TTACTTTTCAACTTCGACTTCGCCACAGACTATTCTCGTAAGGCTCAGAAAGCTTGGTT	300
O	y	301	AGCGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATTA	360
D	b	301	AGCGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATTA	360
O	y	361	GGTCTACATGCTCTAATGGGAACTATGGAATGACAGACAGAGCTCAAAACTATATGTGTA	420
D	b	361	GGTCTACATGCTCTAATGGGAACTATGGAATGACAGACAGAGCTCAAAACTATATGTGTA	420
O	y	421	AGACCTCAATTAATTAAGTTTACCTTCAGTTCAGCTTGCTGGCTGAGATCCTCAGACACC	480
D	b	421	AGACCTCAATTAATTAAGTTTACCTTCAGTTCAGCTTGCTGGCTGAGATCCTCAGACACC	480
O	y	481	AAACCAATATGACCCCTATTTCACATCCAGAGGACGCCAAGACCGCGCAAACTTGGCTT	540
D	b	481	AAACCAATATGACCCCTATTTCACATCCAGAGGACGCCAAGACCGCGCAAACTTGGCTT	540
O	y	541	ATTCGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTAATAC	600
D	b	541	ATTCGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTAATAC	600
O	y	601	ACCAATTACTATGAGCTATACAAGGTCATCAATAGCAAGTAATTACCCTGTTACATGGA	660
D	b	601	ACCAATTACTATGAGCTATACAAGGTCATCAATAGCAAGTAATTACCCTGTTACATGGA	660
O	y	661	TAAATTACCTCAAGGAAGTCATCAATCAAGTTGAAGAAAGAACAGGCTATTAACCTACTAC	720
D	b	661	TAAATTACCTCAAGGAAGTCATCAATCAAGTTGAAGAAAGAACAGGCTATTAACCTACTAC	720
O	y	721	AACGTGGATGATGTCTACACCAATGTAGACCAAGAGCTCAAAAACATCTGGGATAT	780
D	b	721	AACGTGGATGATGTCTACACCAATGTAGACCAAGAGCTCAAAAACATCTGGGATAT	780
O	y	781	TTTCAATATACAGAGCAATACGTTGGCTATCCAGACGATGAATTCGAAATTCGCTTCTACAT	840
D	b	781	TTTCAATATACAGAGCAATACGTTGGCTATCCAGACGATGAATTCGAAATTCGCTTCTACAT	840
O	y	841	TGTTGATGTTTCTAACGCTAAAGTCATTTGCCAGCTAGAGAGACGCGCATCACTCAAGTAA	900
D	b	841	TGTTGATGTTTCTAACGCTAAAGTCATTTGCCAGCTAGAGAGACGCGCATCACTCAAGTAA	900
O	y	901	TGTTTCTCTGGGAATTAACCAAGCAGTAGAACAAACCGGAGCTGGGGATCAACTATGAA	960
D	b	901	TGTTTCTCTGGGAATTAACCAAGCAGTAGAACAAACCGGAGCTGGGGATCAACTATGAA	960
O	y	961	ACCGATCACAGACTATGCTCTGCGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT	1020
D	b	961	ACCGATCACAGACTATGCTCTGCGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT	1020
O	y	1021	CGTTACAGATGAGCCCTATATCACTACCTGGGACAAATACTCCTGTTTATTAACCTGGGATAG	1080
D	b	1021	CGTTACAGATGAGCCCTATATCACTACCTGGGACAAATACTCCTGTTTATTAACCTGGGATAG	1080
O	y	1081	GGGCTACTTTGGCAACATCACTTGCATATACGCCCTGCAACATGCGCAAGAGTCCCAAGC	1140
D	b	1081	GGGCTACTTTGGCAACATCACTTGCATATACGCCCTGCAACATGCGCAAGAGTCCCAAGC	1140

QY	1141	CGTGGAAACTCTAAACAAGGCTGCGACTCAACCGGCGCCACAGACTTCTTAANTGCTTAGG	1200
Db	1141	CGTGGAAACTCTAAACAAGGCTGCGACTCAACCGGCGCCAGACTTCTTAANTGCTTAGG	1200
QY	1201	AATGCACTACCCAACTATCTCACTACTCTCAAAATGCGCAATTTTCAGATGTAACCAACCCGAATGAGA	1260
Db	1201	AATGCACTACCCAACTATCTCACTACTCTCAAAATGCGCAATTTTCAGATGTAACCAACCCGAATGAGA	1260
QY	1261	CAAAAAAATATGAGCAGACAGTAGCTGAAAAAGATGGCTGCTGTACCTGCTCTTGGCAATGG	1320
Db	1261	CAAAAAAATATGAGCAGACAGTAGCTGAAAAAGATGGCTGCTGTACCTGCTCTTGGCAATGG	1320
QY	1321	TGGAACTTACTATTAACCAATGTATATCCATPAAGTGTCTTTAGTATGGAGTGAATA	1380
Db	1321	TGGAACTTACTATTAACCAATGTATATCCATPAAGTGTCTTTAGTATGGAGTGAATA	1380
QY	1381	AGAGTTCTCTTAATGTCGGAACCTGCGCCATGAGAGAAACGACACCTATATGATGACCGA	1440
Db	1381	AGAGTTCTCTTAATGTCGGAACCTGCGCCATGAGAGAAACGACACCTATATGATGACCGA	1440
QY	1441	CATGATGAAAAACAGTCTTGTAGCTATATGGAACAGTGAAGAAATGCTATCTTGGCTGCC	1500
Db	1441	CATGATGAAAAACAGTCTTGTAGCTATATGGAACAGTGAAGAAATGCTATCTTGGCTGCC	1500
QY	1501	TCAGGCTGTATTAACAGACACCTCTAACTATACAGAGAGAGAAATTTGAAAAACCATCAAA	1560
Db	1501	TCAGGCTGTATTAACAGACACCTCTAACTATACAGAGAGAGAAATTTGAAAAACCATCAAA	1560
QY	1561	GACCTCTCAATTTGTAGACACCTGATGAACTTTTGTGGGCTATACGGGTAAATTTTCAT	1620
Db	1561	GACCTCTCAATTTGTAGACACCTGATGAACTTTTGTGGGCTATACGGGTAAATTTTCAT	1620
QY	1621	GCGTGTATGAGACAGGCTATTTCTAACCGTCTGACACACCACTTGTAGGCAATGGCCTTACGGT	1680
Db	1621	GCGTGTATGAGACAGGCTATTTCTAACCGTCTGACACACCACTTGTAGGCAATGGCCTTACGGT	1680
QY	1681	CGCTCGCAAAAGTTTACCGCTCTATGATGAGACTTACTGTCTGAGAGAACATTCAGAAAGA	1740
Db	1681	CGCTCGCAAAAGTTTACCGCTCTATGATGAGACTTACTGTCTGAGAGAACATTCAGAAAGA	1740
QY	1741	TTGGAATTTACACAGAGGGGCTCTACAGAAATGGAATTCGTATTTAAAAATGGTGGCTCG	1800
Db	1741	TTGGAATTTACACAGAGGGGCTCTACAGAAATGGAATTCGTATTTAAAAATGGTGGCTCG	1800
QY	1801	TTCTACGTGGAAGCTACCTGCTCCACAAACAACCCCACTCAACTGAAAGTTTCAAGCTTCATC	1860
Db	1801	TTCTACGTGGAAGCTACCTGCTCCACAAACAACCCCACTCAACTGAAAGTTTCAAGCTTCATC	1860
QY	1861	ATCAGATAGTTCACCTTACAGAGTCTAGCTCAACAGCTCCAGACCAAAATATATAGTACGAC	1920
Db	1861	ATCAGATAGTTCACCTTACAGAGTCTAGCTCAACAGCTCCAGACCAAAATATATAGTACGAC	1920
QY	1921	TACCAATCTTAACATATATATAGCAACAATCAAAATACCAACCCGTATCAACAAAATCGAA	1980
Db	1921	TACCAATCTTAACATATATATAGCAACAATCAAAATACCAACCCGTATCAACAAAATCGAA	1980
QY	1981	TCCTCAACGACACACCA 1999	
Db	1981	TCCTCAACGACACACCA 1999	
RESULT 4			
; Sequence 81, Application US/60029960			
; GENERAL INFORMATION:			
; APPLICANT: Charles Kunsch			
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences			
; NUMBER OF SEQUENCES: 1649			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Human Genome Sciences, Inc.			
; STREET: 9410 Key West Avenue			
; CITY: Rockville			



OY 1681 CCGTCGCAAAAGTTTACCGCTGTATGATGACCTTACCTGTCGTGAAGAGCAATCCAGAAGA 1740  
|||||  
Db 8083 CCGTCGCAAAAGTTTACCGCTGTATGATGACCTTACCTGTCGTGAAGAGCAATCCAGAAGA 8024  
OY 1741 TTGGAATATACAGAGAGGGGCTCTACAGAAATGAGAAATTCGTATTTAAATAATGTCGTG 1800  
|||||  
Db 8023 TTGGAATATACAGAGAGGGGCTCTACAGAAATGAGAAATTCGTATTTAAATAATGTCGTG 7964  
OY 1801 TTCTACGTGGAAGTCACTGCTGTCACAAACACCCCATCACTGAAAGTTCAAGCTCATC 1860  
|||||  
Db 7963 TTCTACGTGGAAGTCACTGCTGTCACAAACACCCCATCACTGAAAGTTCAAGCTCATC 7904  
OY 1861 ATCAATATGTTCACTTCACTGACAGTCAACCACTCCCAACGACAAATAATGATGAC 1920  
|||||  
Db 7903 ATCAATATGTTCACTTCACTGACAGTCAACCACTCCCAACGACAAATAATGATGAC 7844  
OY 1921 TACCAATCTCAATTAATACGCAACAATCAATCAACCCCTGATCAACAATAATGAC 1980  
|||||  
Db 7843 TACCAATCTCAATTAATACGCAACAATCAATCAACCCCTGATCAACAATAATGAC 7784  
OY 1981 TCCTCAACCGACACCA 1999  
|||||  
Db 7783 TCCTCAACCGACACCA 7765

RESULT 5  
US-08-961-527-145/c  
; Sequence 145, Application US/08961527  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 145:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10711 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-961-527-145

Query Match 100.0%; Score 1999; DB 13; Length 10711;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TAAATCTACGACATATAAATCACTCTTGTCTGACTTGGTGTCTGAACGGCGGTCAA 60

Db 9767 TAAATCTACGACATATAAATCACTCTTGTCTGACTTGGTGTCTGAACGGCGGTCAA 9708  
OY 61 TCGCCCAAGTAATGATATTCCACAGATTTGGTTAAGCAATCGTTCTATGAGAACCA 120  
|||||  
Db 9707 TCGCCCAAGTAATGATATTCCACAGATTTGGTTAAGCAATCGTTCTATGAGAACCA 9648  
OY 121 TGGCTTCTTGACACAGGGGGATGATATCCATCCGTATCCGAGGAGCTTCTTGGGCA 180  
|||||  
Db 9647 TGGCTTCTTGACACAGGGGGATGATATCCATCCGTATCCGAGGAGCTTCTTGGGCA 9588  
OY 181 TCGCAAAACCAATTCCTCCCAAGGTGATCACTCCACCCCAAGATTGATTAAGTTGAC 240  
|||||  
Db 9587 TCGCAAAACCAATTCCTCCCAAGGTGATCACTCCACCCCAAGATTGATTAAGTTGAC 9528  
OY 241 TTACTTTTCACTTGACATCTCGACACAGTATTTCTGTAAAGCTCAGAAAGCTTGGTT 300  
|||||  
Db 9527 TTACTTTTCACTTGACATCTCGACACAGTATTTCTGTAAAGCTCAGAAAGCTTGGTT 9468  
OY 301 AGCGATTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTGACCTACTATATAATA 360  
|||||  
Db 9467 AGCGATTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTGACCTACTATATAATA 9408  
OY 361 GGTCTACATGTCTTAATGGGAATGGAATGAGACAGCAGCTCAAACTACTATGTA 420  
|||||  
Db 9407 GGTCTACATGTCTTAATGGGAATGGAATGAGACAGCAGCTCAAACTACTATGTA 9348  
OY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGTGGCTGGAATCCTCAGGACCC 480  
|||||  
Db 9347 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGTGGCTGGAATCCTCAGGACCC 9288  
OY 481 AAACCAATATGACCCCTATTTCATCTCAGAACAGACGCCCAAGACCGGCAAACTTGGTCT 540  
|||||  
Db 9287 AAACCAATATGACCCCTATTTCATCTCAGAACAGACGCCCAAGACCGGCAAACTTGGTCT 9228  
OY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACGTTTGAAGAAAGCATCAATAC 600  
|||||  
Db 9227 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACGTTTGAAGAAAGCATCAATAC 9168  
OY 601 ACCAATTAATGATGACCTCAACAAAGTCTCAATCAGCAAGTAATTACCTGCTTACATGA 660  
|||||  
Db 9167 ACCAATTAATGATGACCTCAACAAAGTCTCAATCAGCAAGTAATTACCTGCTTACATGA 9108  
OY 661 TTAATTAACCTCAAGGAAGTATATCAATCACTTGAAGAAAGAAAGGCTATTAACCTACTCAC 720  
|||||  
Db 9107 TTAATTAACCTCAAGGAAGTATATCAATCACTTGAAGAAAGAAAGGCTATTAACCTACTCAC 9048  
OY 721 AACTGGATGATGTCGTACACAAATGTAGACCAAGAAAGCTCAAAAAATCTGCGGATAT 780  
|||||  
Db 9047 AACTGGATGATGTCGTACACAAATGTAGACCAAGAAAGCTCAAAAAATCTGCGGATAT 8988  
OY 781 TTACAATACAGACGAATACGTTGCTATCCAGACGATGAATTCAGATGCTTACAT 840  
|||||  
Db 8987 TTACAATACAGACGAATACGTTGCTATCCAGACGATGAATTCAGATGCTTACAT 8928  
OY 841 TGTGATGTTTCTAACGGTAAGTCAATGCCAGCTAGAGACAGCGCATCACTCAAGTAA 900  
|||||  
Db 8927 TGTGATGTTTCTAACGGTAAGTCAATGCCAGCTAGAGACAGCGCATCACTCAAGTAA 8868  
OY 901 TGTTCCTTCCGAAATTAACCAAGAGTAGAACAACCGCGATCGGGGATCAATGTGA 960  
|||||  
Db 8867 TGTTCCTTCCGAAATTAACCAAGAGTAGAACAACCGCGATCGGGGATCAATGTGA 8808  
OY 961 ACCGATACAGACTATGCTCTGCTTGAGTAGGCTGTCTAGATTCAACTGCTACTAT 1020  
|||||  
Db 8807 ACCGATACAGACTATGCTCTGCTTGAGTAGGCTGTCTAGATTCAACTGCTACTAT 8748  
OY 1021 CGTTACAGATGAGCCCTATTAACCTACCTGGGACAAATTAATCTCTGTTTAATCTGGATAG 1080  
|||||  
Db 8747 CGTTACAGATGAGCCCTATTAACCTACCTGGGACAAATTAATCTCTGTTTAATCTGGATAG 8688  
OY 1081 GGGTACTTGGCAACATACCTGTGCAATAGCCCGGCAAAATGCGCAAGTCCACAC 1140  
|||||

```
Db 8687 GGGCTACTTTGGCAATCATCTTGCAATACGCCCTGCGACAAATCGGAAATGCCACAC 8628
Qy 1141 CGTGAACACTCTAAACAAGTGGAGCTCAACCGCGCCAAAGCTTTCCTAAATGGCTCTAGG 1200
Db 8627 CGTGAACACTCTAAACAAGTGGAGCTCAACCGCGCCAAAGCTTTCCTAAATGGCTCTAGG 8568
Qy 1201 AATGCACTACCCAGATATCTCTACTCAATGCCATTTTCAAGTAACACACGGAATCAGA 1260
Db 8567 AATGCACTACCCAGATATCTCTACTCAATGCCATTTTCAAGTAACACACGGAATCAGA 8508
Qy 1261 CAAAAAATATGAGACAGTACTGTGAAAAAGATGGCGTGGCTTAACGCTTGTGCAAAATGG 1320
Db 8507 CAAAAAATATGAGACAGTACTGTGAAAAAGATGGCGTGGCTTAACGCTTGTGCAAAATGG 8448
Qy 1321 TGAACCTACTATTAACCAATGTATATCCATTAAGTGTCTTTAGTGTGAGTGAATAA 1380
Db 8447 TGAACCTACTATTAACCAATGTATATCCATTAAGTGTCTTTAGTGTGAGTGAATAA 8388
Qy 1381 AGAGTTCTTAATGTCCGAACTCGTCCATGAGGAAACGACAGCTATATGATGACCGA 1440
Db 8387 AGAGTTCTTAATGTCCGAACTCGTCCATGAGGAAACGACAGCTATATGATGACCGA 8328
Qy 1441 CATATGAAAAACAGTCTTGAATATGAACTGAACTGAAATGCTATCTTGTGGCTGCC 1500
Db 8327 CATATGAAAAACAGTCTTGAATATGAACTGAACTGAAATGCTATCTTGTGGCTGCC 8268
Qy 1501 TCAGAGTGTGTAACAGAACTCTTAATATACAGAGAGGAAATTTGAAAAACCAATCAA 1560
Db 8267 TCAGAGTGTGTAACAGAACTCTTAATATACAGAGAGGAAATTTGAAAAACCAATCAA 8208
Qy 1561 GACCTTCAATTTGTAGACCTGATGAACCTATTTGTGGCTATACGCGTAAATATTCAT 1620
Db 8207 GACCTTCAATTTGTAGACCTGATGAACCTATTTGTGGCTATACGCGTAAATATTCAT 8148
Qy 1621 GGGCTATGGAAGAGGAGTATCTAACCGTTCAGACCACTTGTAGGCAATGGCCTTAAGCT 1680
Db 8147 GGGCTATGGAAGAGGAGTATCTAACCGTTCAGACCACTTGTAGGCAATGGCCTTAAGCT 8088
Qy 1681 CGCTGCCAAAGTTTACGCTCTATGATGACCTTACCTGTCTGAAGGAACATCCAGAA 1740
Db 8087 CGCTGCCAAAGTTTACGCTCTATGATGACCTTACCTGTCTGAAGGAACATCCAGAA 8028
Qy 1741 TTGGAATATACAGAGGAGCTCTACAGAAATGAGAAATTCGTATTTAAAAATGGTCTCG 1800
Db 8027 TTGGAATATACAGAGGAGCTCTACAGAAATGAGAAATTCGTATTTAAAAATGGTCTCG 7968
Qy 1801 TTTCAGTGGAACTCACCTGCTCCACAACAACCCCATCACTGAAGTTCAAGCTCATC 1860
Db 7967 TTTCAGTGGAACTCACCTGCTCCACAACAACCCCATCACTGAAGTTCAAGCTCATC 7908
Qy 1861 ATCAGATAGTTCAACTCACAGTCTAGCTCAACCACTCCAGCAACAATAATATAGTAGAC 1920
Db 7907 ATCAGATAGTTCAACTCACAGTCTAGCTCAACCACTCCAGCAACAATAATATAGTAGAC 7848
Qy 1921 TACCAATCTTACCAATATATACGCAACAATCAATAACACCCCTGATCAACAAATCAGAA 1980
Db 7847 TACCAATCTTACCAATATATACGCAACAATCAATAACACCCCTGATCAACAAATCAGAA 7788
Qy 1981 TCCCAACAGCAGCAACAA 1999
Db 7787 TCCCAACAGCAGCAACCA 7769

RESULT 6
US-09-583-110-1312
; Sequence 1312, Application US/09583110
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583.110
; CURRENT FILING DATE: 2000-05-26
```

```
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 1312
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-1312

Query Match 99.4%; Score 1987.8; DB 22; Length 2160;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAAATCTACGACAAATAAATCAACTCATTTGCTGACTTGGGTTCTGTGACCGCCGCTCAA 60
Db 159 taaatctacgacaaataaatacaactcatctgctgacttgggttctgtgacgcgcgtcaa 218
Qy 61 TGCCCAAGCTAATGATATTTCCACAGATTTGTTAAGGCAATCGTTCTATCGAAGACCA 120
Db 159 taaatctacgacaaataaatacaactcatctgctgacttgggttctgtgacgcgcgtcaa 218
Qy 219 tgccttcttcgacacagaggagattgataccatccgtatccctggagagcttcttcgcaaa 180
Db 219 tgccttcttcgacacagaggagattgataccatccgtatccctggagagcttcttcgcaaa 338
Qy 279 tgccttcttcgacacagaggagattgataccatccgtatccctggagagcttcttcgcaaa 338
Db 279 tgccttcttcgacacagaggagattgataccatccgtatccctggagagcttcttcgcaaa 338
Qy 181 TCTGCAAAAGCAATTCCTCCAGAGTGCATCACTCCACCAACAGTTGATTAAAGTTGAC 240
Db 339 tctgcaaaagcaatctcctccagagtgcatcactccacccaacagttgattaaagttgac 398
Qy 241 TTACTTTTCAACTCGACTTCGACAGACTATTTCTGTGAAGGCTCAGAAAGCTTGGTT 300
Db 399 ttacttttcaactcgcacttcgacagactatcttctgtgaaggctcagaaagcttggtt 458
Qy 301 AGCGATTCAGTTAGAACAAAAGAACCAACCAAGCAAGAAATCTTGACTTATATATAA 360
Db 459 agcgattcagttagaaacaaaagaaacaaacaaagcaagaaatcttgacttataataaa 518
Qy 361 GGTCTACATGCTTAATGGGAATGTGAATGACAGACAGCTCAAAACTTCTATGTGTA 420
Db 519 ggtctacatgcttaatgggaatgtgaatgacagacagctcaaaacttctatgtgtaa 578
Qy 421 AGACCTCAATTAATTAAGTTTACCTGAGTTAGCTTGGCTGGCTGGAATGCTCAGGAC 480
Db 579 agacctcaatataatataagtttacctgagtttagcttggctggctggaatgctcaggac 638
Qy 481 AAACCAATATGACCCCTATTTCATCCAGAGACAGCCCAAGCCGCGAAACTTGCTT 540
Db 639 aaaccaatataatataagtttacctgagtttagcttggctggctggaatgctcaggac 698
Qy 541 ATCTGAATGAAAAATCAAGGCTTACATCTTCTGTAACAGTATGAGAAAGCATATAC 600
Db 699 atctgaatgaaaaatcaaggcttacatcttctgtaaacagtatgagaaagcatatac 758
Qy 601 ACCAATTAATGATGACTACAGAACTCAATTCAGAAAGTAAATACCTGGTTTCATGGA 660
Db 759 accaatataatgactacagaaactcaattcagaaagttaaatacctggtttcattgga 818
Qy 661 TAAATTAATCAAGAGATCATCATCAAGTTGAAGAAAGAAACAGCTTATTAACACTAC 720
Db 819 taattactcaaggaagatcatcatcaagttgaaagaaagaaacagctataaccataccac 878
Qy 721 AACTGGATGATGCTTACACAAATGTAGACCAAGAAAGCTCAAAACATCTGTGGATAT 780
Db 879 aactggatgatgcttaccacaaatgtagaccagaaagctcaaaacatctgtggatatt 938
Qy 781 TTACAATACAGAGCAATACGTTGCTATCCAGAGATGAATGGAAGCGGTTTACCAT 840
Db 939 ttacaatacagacgaataacgcttgcctaccagacgaatgaatggaagcggtttaccat 998
```

[illegible]

```

OY      1981 TCCTCAACGACAGAACCA 1999
Db      2079 taccaatcctaacaataaagaacaaatacaaaccccgatcaacaatacagaa 2138
        2139 taccataaccgagcacacca 2157

RESULT 7
US-09-107-433-1102
; Sequence 1102, Application US/09107433
GENERAL INFORMATION:
APPLICANT: Lynn A Doucet-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
City: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <unknown>
OPERATING SYSTEM: <unknown>
SOFTWARE: <unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8217
INFORMATION FOR SEQ ID NO: 1102:
SEQUENCE CHARACTERISTICS:
LENGTH: 2166 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHEetical: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...2166
SEQUENCE DESCRIPTION: SEQ ID NO: 1102:
US-09-107-433-1102

Query Match          99.4%: Score 1987.8; DB 15; Length 2166;
Best Local Similarity 99.6%: Pred. No. 0;
Matches 1992: Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY      1 TAAATCTACGACATAAATAACAACATCATGTGACTTGGTTCTGAACCGCGCTCAA 60
        |||||||
Db      165 TAAATCTACGCACATAAATAACAACATCATGTGACTTGGTTCTGAACCGCGCTCAA 224
        |||||||
OY      61 TGCCCAAGCTAATGATATTCGCCACAGATTGGTTAAGCAATCGTTTATCGAAGACA 120
        |||||||
```



Db 225 TGCCAAAGCTAATGATATTCACAGATTTTGTTAAGCAATGTTTCTATCGAAGCA 284  
QY 121 TCGCTTCTTGACACAGGGGGATTTGATACATCCGATCCTGGAGCTTTCTTGCCAA 180  
Db 285 TCGCTTCTTGACACAGGGGGATTTGATACATCCGATCCTGGAGCTTTCTTGCCAA 344  
QY 181 TCGCAAGCAATTCCTCCAGAGTGATCACTCCACCAACAGTTTGAATTAAGTTAC 240  
Db 345 TCGCAAGCAATTCCTCCAGAGTGATCACTCCACCAACAGTTTGAATTAAGTTAC 404  
QY 241 TTTACTTTTCACTTCGACTTCGACCAAGACTATTTCTCGTAAGCTCAGAGCTGGTT 300  
Db 405 TTTACTTTTCACTTCGACTTCGACCAAGACTATTTCTCGTAAGCTCAGAGCTGGTT 464  
QY 301 AGGATTCAGTTAGAACAAAGCAACCAAGCAAGAAATTTGAGCCACTATTAATAA 360  
Db 465 AGGATTCAGTTAGAACAAAGCAACCAAGCAAGAAATTTGAGCCACTATTAATAA 524  
QY 361 GGTCTACATGTCTAATGGAACTATGGAATGCAGACAGCTCAAAACTACTATGTAA 420  
Db 525 GGTCTACATGTCTAATGGAACTATGGAATGCAGACAGCTCAAAACTACTATGTAA 584  
QY 421 AGACCTCAATATTTAAGTTTACCTCAATTAGCCTTGCTGGTGAATGGCTCAGCACC 480  
Db 585 AGACCTCAATATTTAAGTTTACCTCAATTAGCCTTGCTGGTGAATGGCTCAGCACC 644  
QY 481 AAACCAATATGACCCCTATTCACATCGAAGAGCAGCCCAAGCCGCAAACTGGCTT 540  
Db 645 AAACCAATATGACCCCTATTCACATCGAAGAGCAGCCCAAGCCGCAAACTGGCTT 704  
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGTGAACAGTATGAGAAAGCACTCAATAC 600  
Db 705 ATCTGAATGAAAAATCAAGGCTACATCTCTGTGAACAGTATGAGAAAGCACTCAATAC 764  
QY 601 ACCAATTAATGAGAGCTACAAAGTCTCAATATGAGCAAGTAATTAACCTCTTACATGA 660  
Db 765 ACCAATTAATGAGAGCTACAAAGTCTCAATATGAGCAAGTAATTAACCTCTTACATGA 824  
QY 661 TAATTAATGAGAGCTACATCAATCAAGTTGAAGAAAGCAAGCTATTAACCTCTTAC 720  
Db 825 TAATTAATGAGAGCTACATCAATCAAGTTGAAGAAAGCAAGCTATTAACCTCTTAC 884  
QY 721 AACTGGATGAGATGTCTACCAAAATGTAGCAAGAAAGCTCAAAACATCTGTGGATAT 780  
Db 885 AACTGGATGAGATGTCTACCAAAATGTAGCAAGAAAGCTCAAAACATCTGTGGATAT 944  
QY 781 TTTCAATACAGAGATACGTTGCTATCCAGACGATGAATTTGCAAGTCCCTCTACAT 840  
Db 945 TTTCAATACAGAGATACGTTGCTATCCAGACGATGAATTTGCAAGTCCCTCTACAT 1004  
QY 841 TGTGATGTTTAAAGGTTAAAGTATGCTGCGAGTAGAGACAGCCATCAGTCAAGTAA 900  
Db 1005 TGTGATGTTTAAAGGTTAAAGTATGCTGCGAGTAGAGACAGCCATCAGTCAAGTAA 1064  
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAAACAAACCGGATGGGATCAACTAGAA 960  
Db 1065 TGTTCCTTCGGAATTAACCAAGCAGTAGAAACAAACCGGATGGGATCAACTAGAA 1124  
QY 961 ACCGATACAGACTATGCTCCTGCTGGAGTACGCTGTACGATTCACATCTCTACTAT 1020  
Db 1125 ACCGATACAGACTATGCTCCTGCTGGAGTACGCTGTACGATTCACATCTCTACTAT 1184  
QY 1021 CGTTACAGATGAGCCCTATTAATACCTGGAGCAAAATACTCGTTTATTAATGGGATAG 1080  
Db 1185 CGTTACAGATGAGCCCTATTAATACCTGGAGCAAAATACTCGTTTATTAATGGGATAG 1244  
QY 1081 GGGCTACTTTGGCAACATCCTTGCAATAGCCCTGCAACAAATCGGAAACGCTCCAGC 1140  
Db 1245 GGGCTACTTTGGCAACATCCTTGCAATAGCCCTGCAACAAATCGGAAACGCTCCAGC 1304  
QY 1141 CGTGAAGAACTTAAGAAAGTGGAGTCAACCCGCGCAGAGCTTTCTTAATGTCTAGG 1200  
Db 1305 CGTGAAGAACTTAAGAAAGTGGAGTCAACCCGCGCAGAGCTTTCTTAATGTCTAGG 1364

QY 1201 AATGACTACCAAGTATTCCTACTACTCAATGGCATTTCAAGTAACACACCAATCAGA 1260  
Db 1365 AATGACTACCAAGTATTCCTACTACTACTCAATGGCATTTCAAGTAACACACCAATCAGA 1424  
QY 1261 CAAAAAATATGAGCAAGTAGTAAGAAAGATGCTGCTGCTTACGCTGCTTTCGAAATGG 1320  
Db 1425 CAAAAAATATGAGCAAGTAGTAAGAAAGATGCTGCTGCTTACGCTGCTTTCGAAATGG 1484  
QY 1321 TGGAACTTACTATTAACCAATGTATATCCATTAAGTCGTCTTATAGATGGGAGTCAAA 1380  
Db 1485 TGGAACTTACTATTAACCAATGTATATCCATTAAGTCGTCTTATAGATGGGAGTCAAA 1544  
QY 1381 AGAGTTCTCTAATGTGGAAGCTGCTGCAATGAGAAAGCAACGCACTATTAATGAGCA 1440  
Db 1545 AGAGTTCTCTAATGTGGAAGCTGCTGCAATGAGAAAGCAACGCACTATTAATGAGCA 1604  
QY 1441 CATGATGAAAAAGCTGTTGACTATGGAAGTGAAGCAAAATGCTATCTTGCTGCC 1500  
Db 1605 CATGATGAAAAAGCTGTTGACTATGGAAGTGAAGCAAAATGCTATCTTGCTGCC 1664  
QY 1501 TCAAGCTGTTAAACAGAAAGCTCTAATTAATACAGAGAGGAAATTTGAACCAATCAA 1560  
Db 1665 TCAAGCTGTTAAACAGAAAGCTCTAATTAATACAGAGAGGAAATTTGAACCAATCAA 1724  
QY 1561 GACCTGCAATTTGTAGCACTGATGAACTATTTGCTGGGTATACGGTAAATTAATCAAT 1620  
Db 1725 GACCTGCAATTTGTAGCACTGATGAACTATTTGCTGGGTATACGGTAAATTAATCAAT 1784  
QY 1621 GGCTGATGAGAGGCTATTTCAACGCTGAGACCACTTTGAGCAATGGCTTACGCT 1680  
Db 1785 GGCTGATGAGAGGCTATTTCAACGCTGAGACCACTTTGAGCAATGGCTTACGCT 1844  
QY 1681 CGCTGCCAAAGTTTACCGCTATGATGAACTGCTGCTGGAAGAAAGCAATTCAGAGA 1740  
Db 1845 CGCTGCCAAAGTTTACCGCTATGATGAACTGCTGCTGGAAGAAAGCAATTCAGAGA 1804  
QY 1741 TTGGAATATACAGAGGAGGCTCTACAGAAATGAGAAATTCGTTTAAAAATGGTCTCG 1800  
Db 1905 TTGGAATATACAGAGGAGGCTCTACAGAAATGAGAAATTCGTTTAAAAATGGTCTCG 1964  
QY 1801 TTCTAGTGAAGTCACTGCTGCTCCACACCAACCCCATCACTGAAATGTTCAAGCTCATC 1860  
Db 1965 TTCTAGTGAAGTCACTGCTGCTCCACACCAACCCCATCACTGAAATGTTCAAGCTCATC 2024  
QY 1861 ATCAGATGTTCAAGTCTCAGCTAGCTGCTCAACCACTCCAGCAAAATTAATGTACGAC 1920  
Db 2025 ATCAGATGTTCAAGTCTCAGCTAGCTGCTCAACCACTCCAGCAAAATTAATGTACGAC 2084  
QY 1921 TACCAATTCCTAACAATTAATACGCAACATCAATTAACCCCTGATCAACAAATCAGAA 1980  
Db 2085 TACCAATTCCTAACAATTAATACGCAACATCAATTAACCCCTGATCAACAAATCAGAA 2144  
QY 1981 TCCTCAACGAGCAACCA 1999  
Db 2145 TCCTCAACGAGCAACCA 2163

RESULT 8  
PCT-US02-03987-9325  
; Sequence 9325, Application PC/TUS0203987  
; GENERAL INFORMATION:  
; APPLICANT: Elitza Pharmaceuticals, Inc.  
; TITLE OF INVENTION: Methods for identifying the target of a compound which inhibit  
; FILE REFERENCE: ELITRA.028VPC  
; CURRENT APPLICATION NUMBER: PCT/US02/03987  
; PRIOR FILING DATE: 2002-02-02  
; PRIOR APPLICATION NUMBER: 60/267,636  
; NUMBER OF SEQ ID NOS: 15811  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9325



LENGTH: 2160  
TYPE: DNA  
ORGANISM: Streptococcus pneumoniae  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(2160)  
PCT-US02-03987-9325

Query Match 99.1%; Score 1981.4; DB 1; Length 2160;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTCGCAATTAATCACTCATGTGCTGGGTGCGGCGGTCA 60  
Db 159 TAAATCTCGCAATTAATCACTCATGTGCTGGGTGCGGCGGTCA 218  
QY 61 TCCCAAGCTAATGATATCCACAGATTGGTTAGGCAATCGTTCTATCGAAGCA 120  
Db 219 TCCCAAGCTAATGATATCCACAGATTGGTTAGGCAATCGTTCTATCGAAGCA 278  
QY 121 TCGCTTCTTGACACACAGGGGATGATACATCGTATCTGGAGCTTTCTGCGCA 180  
Db 279 TCGCTTCTTGACACACAGGGGATGATACATCGTATCTGGAGCTTTCTGCGCA 338  
QY 181 TCGCAAGCAATTCCTCCAGGTGATCACTCTCACCCCAAGTGTGAAGTTGAC 240  
Db 339 TCGCAAGCAATTCCTCCAGGTGATCACTCTCACCCCAAGTGTGAAGTTGAC 398  
QY 241 TTACTTTTCACTGACTTCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTT 300  
Db 399 TTACTTTTCACTGACTTCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTT 458  
QY 301 ACGGATTCAGTTAGAAACAAAACAAACAGCAAGAAATCTTGACTTATATAATA 360  
Db 459 ACGGATTCAGTTAGAAACAAAACAAACAGCAAGAAATCTTGACTTATATAATA 518  
QY 361 GGTCTCATGTCTAATGGGAACATGATGACAGAGAGCTCAAACTACTANTGTAA 420  
Db 519 GGTCTCATGTCTAATGGGAACATGATGACAGAGAGCTCAAACTACTANTGTAA 578  
QY 421 AAGCTCAATTAATTAAGTTTACCTCAGTTAGGCTTGCTGGATGCTTCAGGAC 480  
Db 579 AAGCTCAATTAATTAAGTTTACCTCAGTTAGGCTTGCTGGATGCTTCAGGAC 638  
QY 481 AAGCAATATGACCCCTATTACATTCAGAAAGCAGCCCAAGAGCCGCAAACTTGTCT 540  
Db 639 AAGCAATATGACCCCTATTACATTCAGAAAGCAGCCCAAGAGCCGCAAACTTGTCT 698  
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGTGAACGATAGAAAGCATCAATAC 600  
Db 699 ATCTGAAATGAAAAATCAAGGCTACATCTCTGTGAACGATAGAAAGCATCAATAC 758  
QY 601 ACCAATTAATGATGACTCAAAAGTCTCAATAGCAAGTAATTAATCTGCTTACATGA 660  
Db 759 ACCAATTAATGATGACTCAAAAGTCTCAATAGCAAGTAATTAATCTGCTTACATGA 818  
QY 661 TAATTAAGTCAAGAGATCATCAATCAAGTTGAAGAAAGAGGCTATTAACCTTACTAC 720  
Db 819 TAATTAAGTCAAGAGATCATCAATCAAGTTGAAGAAAGAGGCTATTAACCTTACTAC 878  
QY 721 AACTGGGATGATGTGTACCAAAATGTAGACCAAGAGCTCAAAACATCTGGGATAT 780  
Db 879 AACTGGGATGATGTGTACCAAAATGTAGACCAAGAGCTCAAAACATCTGGGATAT 938  
QY 781 TTACATTAAGAGAGATAGCTTCTATCCAGAGATGAATTCAGATGCTGCTTACAT 840  
Db 939 TTACATTAAGAGAGATAGCTTCTATCCAGAGATGAATTCAGATGCTGCTTACAT 998  
QY 841 TGTGATGTTTCTAACGGTAAATCAATGGCCAGCTAGAGAGACGCCATCATGCTAAATGA 900  
Db 999 TGTGATGTTTCTAACGGTAAATCAATGGCCAGCTAGAGAGACGCCATCATGCTAAATGA 1058

QY 901 TGTTCCTTCGGATTAACAGAGTAGAACAACCGAGCTGGGATCACTATGA 960  
Db 1059 TGTTCCTTCGGATTAACAGAGTAGAACAACCGAGCTGGGATCACTATGA 1118  
QY 961 ACCGATACAGACTATGCTCTGCTTGAAGTACGGTGTCTACGATTCATCTAT 1020  
Db 1119 ACCGATACAGACTATGCTCTGCTTGAAGTACGGTGTCTACGATTCATCTAT 1178  
QY 1021 CGTTACAGTAGAGCCCTATTAACCTGGGACAAATTAATCTCCGTATTAACGGGATAG 1080  
Db 1179 CGTTACAGTAGAGCCCTATTAACCTGGGACAAATTAATCTCCGTATTAACGGGATAG 1238  
QY 1081 GGGCTACTTGGCAACATCACTTGCATATACGGCTGCAACAATCGCAAACTGCTCAC 1140  
Db 1239 GGGCTACTTGGCAACATCACTTGCATATACGGCTGCAACAATCGCAAACTGCTCAC 1298  
QY 1141 CGTGAACCTTAACAAAGGTGAGCTCAACCGCGCAAGACTTCTTAATGCTGTAG 1200  
Db 1299 CGTGAACCTTAACAAAGGTGAGCTCAACCGCGCAAGACTTCTTAATGCTGTAG 1358  
QY 1201 AATCGACTACCCAGATTAATCACTACATTAATGCAATTAAGTAACACGCAATCAG 1260  
Db 1359 AATCGACTACCCAGATTAATCACTACATTAATGCAATTAAGTAACACGCAATCAG 1418  
QY 1261 CAAAAATATGAGCAAGTAGTAAGAAAGATGCTGCTTACGCTTGCCTTGAATG 1320  
Db 1419 CAAAAATATGAGCAAGTAGTAAGAAAGATGCTGCTTACGCTTGCCTTGAATG 1478  
QY 1321 TGGAACTTACTATTAACCAATGATATTCATAAAGTCTTATAGTATGAGTCAAAA 1380  
Db 1479 TGGAACTTACTATTAACCAATGATATTCATAAAGTCTTATAGTATGAGTCAAAA 1538  
QY 1381 AGAGTTCTTAATGTTGGAACTGTGTCCATGAAGAAAGACGACCTTATATATGATGACCA 1440  
Db 1539 AGAGTTCTTAATGTTGGAACTGTGTCCATGAAGAAAGACGACCTTATATATGATGACCA 1598  
QY 1441 CATGATGAACAAAGCTTGTGACTTATGAACCTGACCAAAATGCTTATCTGCTGGCTCC 1500  
Db 1599 CATGATGAACAAAGCTTGTGACTTATGAACCTGACCAAAATGCTTATCTGCTGGCTCC 1658  
QY 1501 TCAAGCTGTGTAACAGAGAACTCTTAATATACAGACAGAGAAATTAAGAAACCAATCA 1560  
Db 1659 TCAAGCTGTGTAACAGAGAACTCTTAATATACAGACAGAGAAATTAAGAAACCAATCA 1718  
QY 1561 GACCTCTCAATTTGTAGACCTGATGATGATTTGCTGGCTATACGCTTAATATTCAT 1620  
Db 1719 GACCTCTCAATTTGTAGACCTGATGATGATTTGCTGGCTATACGCTTAATATTCAT 1778  
QY 1621 GGCTGTATGAGACAGGCTATTAACCGTCTGACACCACTGTAGGCAATGAGCTTACGGT 1680  
Db 1779 GGCTGTATGAGACAGGCTATTAACCGTCTGACACCACTGTAGGCAATGAGCTTACGGT 1838  
QY 1681 CGCTGCCAAAGTTTACCGCTTATGATGATGATCTGCTGGAAGAGCAATTCAGAA 1740  
Db 1839 CGCTGCCAAAGTTTACCGCTTATGATGATGATCTGCTGGAAGAGCAATTCAGAA 1898  
QY 1741 TTGGAATATACAGAGAGGCTCTACAGAAATGGAATTCGTTTAAAAATGCTGCTG 1800  
Db 1899 TTGGAATATACAGAGAGGCTCTACAGAAATGGAATTCGTTTAAAAATGCTGCTG 1958  
QY 1801 TTCTAGTGAACCTCACTGCTCCACACCAACCCCATCACTCAAGTAAGTTCAAGCTCAT 1860  
Db 1959 TTCTAGTGAACCTCACTGCTCCACACCAACCCCATCACTCAAGTAAGTTCAAGCTCAT 2018  
QY 1861 ATCAGATAGTTTCAACTTACAGTCTAGCTCAACACTCAAGACAAATATAGTAGAC 1920  
Db 2019 ATCAGATAGTTTCAACTTACAGTCTAGCTCAACACTCAAGACAAATATAGTAGAC 2078  
QY 1921 TACCAATCCCTAACAAATATACGCAACATCAATTAACACCCCTGATCAACAAATACGA 1980  
Db 2079 TACCAATCCCTAACAAATATACGCAACATCAATTAACACCCCTGATCAACAAATACGA 2138  
QY 1981 TCCTCAACGACACACCA 1999

Db 2139 tcctcaaccagcacaacca 2157

## RESULT 9

```

1 Sequence 9325, Application US/09615242
2 GENERAL INFORMATION:
3 APPLICANT: Haselbeck, Robert
4 APPLICANT: Ohlsen, Karl L.
5 APPLICANT: Zysken, Judith W.
6 APPLICANT: Wall, Daniel
7 APPLICANT: Trawick, John D.
8 APPLICANT: Carr, Grant J.
9 APPLICANT: Yamamoto, Robert T.
10 APPLICANT: Xu, H. Howard
11 TITLE OF INVENTION: Identification of Essential Genes in
12 TITLE OF INVENTION: Prokaryotes
13 FILE REFERENCE: ELITR-011a
14 CURRENT APPLICATION NUMBER: US/09/615,242
15 CURRENT FILING DATE: 2001-03-21
16 PRIOR APPLICATION NUMBER: 60/191,078
17 PRIOR FILING DATE: 2000-03-21
18 PRIOR APPLICATION NUMBER: 60/206,848
19 PRIOR FILING DATE: 2000-05-23
20 PRIOR APPLICATION NUMBER: 60/207,757
21 PRIOR FILING DATE: 2000-05-26
22 PRIOR APPLICATION NUMBER: 60/242,578
23 PRIOR FILING DATE: 2000-10-23
24 PRIOR APPLICATION NUMBER: 60/253,625
25 PRIOR FILING DATE: 2000-11-27
26 PRIOR APPLICATION NUMBER: 60/257,931
27 PRIOR FILING DATE: 2000-12-22
28 PRIOR APPLICATION NUMBER: 60/269,308
29 PRIOR FILING DATE: 2001-02-16
30 NUMBER OF SEQ. ID NOS: 14,110
31 SOFTWARE: FastSeq for Windows Version 4.0
32 SEQ ID NO 9325
33 LENGTH: 2160
34 TYPE: DNA
35 ORGANISM: Streptococcus pneumoniae
36 FEATURE:
37 NAME/KEY: CDS
38 LOCATION: (1)..(2160)
39 US-09-815-242-9325

```

Query Match	99.18;	Score 1981.4;	DB 31;	Length 2160;
Best Local Similarity	99.48;	Pred. No. 0;		
Matches 1988; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0;

QY	1	TAAATCTAGCACAATAAAAAATCACTCATTTGCTGACTTGGGTTCTGAAACGCCGCGTCAA	60
Db	159	taaatctacgacaaataaaaaatacaactcatctgacttggttctctgaaagccgagctcaa	218
QY	61	TGCCAAGCTAATATATATTTCACACAGATTGGTTAAAGCAATCGTTCTATACGAAGCCA	120
Db	219	tgcacaagctaatatatttccacaaagatttggtaagaagcaatgcttctctcgaagcca	278
QY	121	TGCGTTCTTCGACCCACAGAGGGGATTTGATACCATCCGTAATCCGTGGAGCTTTCTGGCAA	180
Db	279	tgcgttcttcgacacagaggatattgtatcacatccgatactcctggagcttcttcgcaa	338
QY	181	TCTGCAAAAGCAATTCCTCCACAGGTGATCACTCTACCCACACAGTTGATTAAATTGAC	240
Db	339	tctgcaaaagcaatccctcccaaggtgatcagctctcacatcaacagttgattaagtgcac	398
QY	241	TTACATTTCACACTCGACTTCGACGACGACTATTTCTGTAAAGCTCAGGAAGTTGGTT	300
Db	399	ttacttttcaacttcgacttcgcgcacgactattctccgtaaggtctaaaggaagcttggtt	458
QY	301	AGGATTTCAGTTTAGAACAAAAGACAAACAGCAAGAAATCTTGACCTATCATTTAAATAA	360

Dd	459	agagatcagcttagaacaacaaagcaaccaagaagaatccttgacctactatataataa	518
Qy	361	ggtctacacatgcttaattatggagactatggaatgcagacagcagctcaaaaactactatgctaa	420
Dd	519	ggtctacatgcttaatigggaactatggaatgcagacagcagctcaaaactactatgctaa	578
Qy	421	agacctcaattattttaaagtttacctcagtttagccttgctggaaatgctcagggcacc	480
Dd	579	agacctcaataatctaagctttacctcagtttagccttgctggaaatgctcagggcacc	638
Qy	481	aaaccaaattatgacccttatttcacatccagacagcggcgaagccggcgaatttgctgtt	540
Dd	639	aaaccaaataagacccttcaatccacccaagaagaagcccaagcgcgcgaacttggctt	698
Qy	541	attctgaattgaaaaatcagaagctacattctctgctgaacagtatgagaaaagcagtcattac	600
Dd	699	attctgaattgaaaaatcagaagctacatctctgctgtaacagtatgagaaagcagtcattac	758
Qy	601	accaattactatgctcagctacaaaagctcctaattgcacaaattttacctgctttacattga	660
Dd	759	accaattactatgctcagctacaaaagctcctaattgcacaaattttacctgctttacattga	818
Qy	661	taatttacctcagaagaaactcattcattcaatcagaattggaagaaagaaacagctttaacctctac	720
Dd	819	taatttacctcagaagaaagctcattcattcaatcagaattggaagaaagaaacagctttaacctctac	878
Qy	721	aaactggatggatgtcttcacacaaatgatagacacaaagctcaaaaacattctgtggatatt	780
Dd	879	aaactggatggatgtcttcacacaaatgatagacacaaagctcaaaaacattctgtggatatt	938
Qy	781	tttacaattacagacgaattatgcttgcttatccacagatgaattgcaatgctgctttaccatt	840
Dd	939	tttacaattacagacgaattatgcttgcttatccacagatgaattgcaatgctgctttaccatt	998
Qy	841	tgttgatgcttttctaaagctttaaagctaaagctaaagctcctgaagacagccatcagtcagtaa	900
Dd	999	tgttgatgcttttctaaagctttaaagctaaagctcctgaagacagccatcagtcagtaa	1058
Qy	901	tgttttcccttcggagtttttaaccacagcagtagaamaaaccgagctggggatcacaattgaa	960
Dd	1059	tgttttcccttcggagtttttaaccacagcagtagaamaaaccgagctggggatcacaattgaa	1118
Qy	961	accgattcacagacttatgctcctgctgcttggaatgagctgattcagattcactgctactatt	1020
Dd	1119	accgattcacagacttatgctcctgctgcttggaatgagctgattcagattcactgctactatt	1178
Qy	1021	cgtttcacagatgagacccttattatcattaccctgggacaaattactcctgttttttacctgggattag	1080
Dd	1179	cgtttcacagatgagacccttattatcattaccctgggacaaattactcctgttttttacctgggattag	1238
Qy	1081	ggggctacgttttggcacaacttcacttggcaattacgcccctgcacacantggcgaacagctccagc	1140
Dd	1239	ggggctacgttttggcacaacttcacttggcaattacgcccctgcacacantggcgaacagctccagc	1298
Qy	1141	cgtggaaaactctttaaacaaggctcggactcaaacgcccgaagactttcttaattgcttagg	1200
Dd	1299	cgtggaaaactctttaaacaaggctcggactcaaacgcccgaagactttcttaattgcttagg	1358
Qy	1201	aattgcactaccacagattttcactactcctaattgcattttaaataacacacaggaattcaga	1260
Dd	1359	aattgcactaccacagattttcactactcctaattgcattttaaataacacacaggaattcaga	1418
Qy	1261	caaaaaaattgagacagctgtagaanaagatggctgcttgcgtggcctttggcaaatg	1320
Dd	1419	caaaaaaattgagacagctgtagaanaagatggctgcttgcgtggcctttggcaaatg	1478
Qy	1321	tggaaacttactattaaacaaatgtatttccattaaagctgctttacttgatggagatgaaa	1380
Dd	1479	tggaaacttactattaaacaaatgtatttccattaaagctgctttacttgatggagatgaaa	1538
Qy	1381	agagcttcttatgtcggaactcgtgcccattgaagaacacagacgcttatgatgaccca	1440
Dd	1539	agagcttcttatgtcggaactcgtgcccattgaagaacacagacgcttatgatgaccca	1598

```
QY 1441 CATGATGAAAACAGTCTTACTTATGAGTGGAGAAATGCTATCTTGCTTGCTCC 1500
      |||
Db 1599 catgatgaaaaacgctcttgaftaagaaactggaacgaatgctcctgctgctccc 1658
QY 1501 TCAGGCTGGTAAACAGAACTCTACTATACAGAGAGAAATGAAACCATCAAA 1560
      |||
Db 1659 tcaggcttgtaaaacaggaacctctaacctacagcaggaatgaaacacatcaa 1718
QY 1561 GACCTTCAATTTGTAGCACTGATGAACTATTTGCTGCTATACGCGTAATATTCAAT 1620
      |||
Db 1719 gacctctcaattttagcaccgtagtaactatttgcctgctaaacgctaaatcat 1778
QY 1621 GGCTGATGAGAGGCTATTTTACCGCTGACACCACTGTGAGCAATGGCTTGCTGCT 1680
      |||
Db 1779 ggcctgataagagcctatctcaaccgctctgacacacatctgagcaatcgctta 1838
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGGAAGCAATCCAGAAA 1740
      |||
Db 1839 cgctgcaaaagtctaccgctctatgtagcctaccctgcttgaaaggaatccaga 1898
QY 1741 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCGATTTAAAAATGCTGCTG 1800
      |||
Db 1899 ttggaatatacagaagggctctacagaatgagaattcgtaatttaaaatgctgctg 1958
QY 1801 TTTCTAGTGAAGTCTACCTGCTCTCCACAAACCCCATGCACTGGAAGTTCAAGCTCATC 1860
      |||
Db 1959 ttctaagtgagctcaccgctctccacaacaccccatcaactgaagttcaagctcatt 2018
QY 1861 ATCAGATGATTAACACTTCAAGTCTGCTCAACCACTCCAGACAAATTAATAGTACGAC 1920
      |||
Db 2019 atcagatgattcaacttcaactcagctcagctcaacacactccaagcacaataatagtcagac 2078
QY 1921 TACCAATCTTAACATATATAGCAACATCAATATACACCCCTGATCAACAAATATCAGAA 1980
      |||
Db 2079 taaccaatcctaacaataatagcaacaatcaataacacccctgatacaacaatcagaa 2138
QY 1981 TCTCTAACACGACGACACCA 1999
      |||
Db 2139 tctctaacacgacgacacca 2157

RESULT 10
US-10-072-851-9325
; Sequence 9325, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 9325
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2160)
US-10-072-851-9325

Query Match          99.1%; Score 1981.4; DB 37; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATAAATCAACTATGCTGACTTGCTGCTTGAACGCCGCTCAA 60
      |||
Db 159 taaatctacgacaaataaaatcaactatgctgcttggcttgcgaacgcgcgtcaa 218
QY 61 TGCCCAAGCTAATGATATTCCACACATTTGTTAAGGCAATCGTTTATGCAACCA 120
      |||
Db 219 tgcccaagctaataatgattccacacatttggtaaggaatgcttctatcgaagacca 278
QY 121 TGCTTCTTGACACACAGGGGGATTGATACATCCGATCCGTGAGAGCTTTTGGCGAA 180
      |||
Db 279 tgccttcttgacacacaggggattgataccacgtaaccgtagggagcttcttcgcgcaa 338
QY 181 TCTGCAAAAGCAATTCCTCCAGAGGTGATCACTCTACCCCAACAGTTGATTAACTGAC 240
      |||
Db 339 tctgcaaaagcaattcctcccaaggtgattcagctctcactcaacagttgattaaagttgac 398
QY 241 TTTACTTTTCACTTGACCTTCCGACGACATTTTCTGTAAGGCTTCAGAGAGCTTGT 300
      |||
Db 399 ttactttcaacttcgacttcgacagactattctcgtlaaggtcgaagaagcttgctt 458
QY 301 AGCGATTCAAGTTGAGCAAAAAGCAACCAAGCAAGAAATCTTGACTATTAATATA 360
      |||
Db 459 agcgattcagtttagaacaacaaacaaacaaagaatcttgcactactataataaa 518
QY 361 GGCTCATGCTTAATGGAAGCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
      |||
Db 519 ggtctcatgcttaatggaagcaatgagatgagagagagagagagagagagagagagag 578
QY 421 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGGCTTGCTGCTGGAATGCTTCAGGACCC 480
      |||
Db 579 agacctcaatataatgatttaagttactcagttagcttgctgctgagatgctcgaagcacc 638
QY 481 AAACCAATATGACCCCTATTTCAATCAGACAGAGCCCAAGACCCGCGAACTTGCTCT 540
      |||
Db 639 aaaccaatataagccctattcacaaccagaagagcccaagagccggaacttgctt 698
QY 541 AACTGAAATGAAAAATCAAGGCTACATCTGCTGTAACAGTATGAGAAAGCAGTCAATAC 600
      |||
Db 699 aactgaaatgaaaaatcaagagctacactctctgctgaaacgtatgaaagcagtcacac 758
QY 601 ACCAATTACTGATGAGCTACAAAGTCTCAAAATCAGCAAGTAAATTACCTGCTTACATGA 660
      |||
Db 759 accaattactgattggtctacaagctcacaatcaagaagtaattaccctctacatgaa 818
QY 661 TAATTACCTCAAGAGAGTATCAATCAAGTTGAAGAAAGAAACAGGTATTAACCTATCTAC 720
      |||
Db 819 taattaccccaaggaaggtatcaatcaagttgaagaagaacaggtctataactactcac 878
QY 721 AACTGGATGATGCTGTACAAATATGAGCAAGAGCTCAAAAACATCTGCGGATAT 780
      |||
Db 879 aactggatgattgctgtacaaatattgagacaaaggtcacaacacatctgtggatatt 938
QY 781 TTACATACAGAGCAATAGCTTGCTTCCATACAGAGATGATGCAAGTGCCTTCAACAT 840
      |||
Db 939 ttacaatacagagcaatagcttgcttccatccagagctgattgcaagtcgcttccacat 998
QY 841 TGTGATGTTTCTACAGGTAAGTCAATGCTCCAGCTAGAGACAGCCATCAATGCAAGTAA 900
      |||
Db 999 tgtgatgttcttcaaggttaaggtcattgcccagcttagagacagcattcagtcaggttaa 1058
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTGAAGAAACAAACCCGAGCTGGGATCAACTATGAA 960
      |||
Db 1059 tgttcccttcggaattaaccagcagtagaacaacacgcgactgggaggtcaactatgaa 1118
```

QY 961 ACCGATACAGACTAGCTCTGCTTGGAGTAGCGTGTCTACGATTCAGTCTGCTACTAT 1020  
 DB 1119 accgatcaagactatgctcctccttgagtagcgtgtctacgagtcacgtccacact 1178  
 QY 1021 CGTTACGATGAGCCCTTAATTAATACCTGGGCAAAATACTCCTGTTTATTAATGGGATG 1080  
 DB 1179 cgttcagatgagcccttaactacccttggaacaataccctgttataacttggatag 1238  
 QY 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCTTGCAACAATGGCGAAAGCTCCAGC 1140  
 DB 1239 gggctactttggcaacaatcaccttgcaatagccttgcaacaatcggaagctccagc 1298  
 QY 1141 CGTGGAAACTTAACAAGTGTGAGTCAACCGCGCAAGACTTCTTAATGCTTAGG 1200  
 DB 1299 cgtggaacttaacaagtgtagctcaacgcgcgaagacttccctaatactgctcgg 1358  
 QY 1201 AATGCACTACCAATATTACTACTCAAAATGCCATTTCAATACACACAGCATGACA 1260  
 DB 1359 aatgcaactaccaatattactactcaaaatgccaattcaagtaacacaaccgaatcaga 1418  
 QY 1261 CAAAAAATATGAGAGCAAGTAGTGAAGAAGATGCTGCTTACGCTTGCTTGGCAATAG 1320  
 DB 1419 caaaaaeatatgagcaagtaggaaaaagatggtcgtcttaccgcttgcgaatagg 1478  
 QY 1321 TGGAACTTAATAACCAATGTATATCCATTAAGTCTTCTTAAATGGAGTGAAGA 1380  
 DB 1479 tggaaacttataaaccgaatglatatccataaagtcgtctttagtggtaggaa 1538  
 QY 1381 AGAGTTCTTAATGTCGAACCTGCGCATATGAGAAAGACAGCAACGCTTATGATACCGA 1440  
 DB 1539 agagttccttaatgtcgaacctgcgcaatgaagaacagacactatagatgacga 1598  
 QY 1441 CATGATGAAAAACAGTCTTGAATTAATGAGTGAAGAAATGCTTGTGCTGCTCC 1500  
 DB 1599 catgatgaaaaacagtcttgagtagtgaacgtggaagaatgctctcttgcgtcc 1658  
 QY 1501 TCAGGCTGTAAACAGAGAACCTTAATATACAGAGAGAAATTTGAACACCATCAA 1560  
 DB 1659 tcaggctgttaaaacaggaaccttaactatacagaagagaaatgaaaaccacataa 1718  
 QY 1561 GACCTCTCAATTTTACAGACCTGATGAGACTATTTGCTGCTATACGCTTAATATCAAT 1620  
 DB 1719 gacctctcaatlttgagacctgagaaactatltgctgtatacgcglaaatatccaat 1778  
 QY 1621 GGCTGTATGACAGGCTATTTAAACGCTGTGACACACTTGTAGGACATGGGCTTAGG 1680  
 DB 1779 ggcgtgtatgacagagctatlttaaccgcttgcgaacactttaggcaatggtcctacggt 1838  
 QY 1681 CGCTGCCAAAGTTTACGCTCTATGATGACCTTACCTGTCTGAGAGAGCAATCCAGAGA 1740  
 DB 1839 cgctgccaaagtttacgctctatgatacctactgcttgaagaagcaatccagagga 1898  
 QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTTGATTTAAATGGTCTCG 1800  
 DB 1899 ttggaatataccagagggctcttaagaatcggagaaatltgcttataaataatggtccg 1958  
 QY 1801 TTTCAGTGAAGTCACTGCTGTCCACAACAACCCCATCACTGAAATTTAAGCTCATC 1860  
 DB 1959 ttccaagtgagagctcactgtctcaacaaccccccaactcaactgaagttcaagctc 2018  
 QY 1861 ATCAGATAGTTCAACTTACAGTCTAGCTCAACCACTGCAAGCACAATAATAGTAGAC 1920  
 DB 2019 atcagatagttcaacttcaactgactagctcaacacactcaagcaacaataatagtagac 2078  
 QY 1921 TACCAATCTTAACAATATAGCAACATCAAAATCAACCCCTGTATCAACAATAATCAGAA 1980  
 DB 2079 taccaatcttaacaataatagcaacatcaatacaacccccctgaltcaacaataatcagaa 2138  
 QY 1981 TCCTCAACGACGACACCA 1999  
 DB 2139 tcctcaacgacgacacaa 2157

RESULT 11  
 US-60-061-998-596/c  
 : Sequence 596, Application us/60061998  
 : GENERAL INFORMATION:  
 : APPLICANT: LAGACE, ROBERT E.  
 : APPLICANT: COREY, NEIL C.  
 : APPLICANT: RUSSO, FRANK D.  
 : APPLICANT: HANN, AMY L.  
 : APPLICANT: HEATH, JOE D.  
 : APPLICANT: FINNEY, GREGORY L.  
 : TITLE OF INVENTION: NOCLOBOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE  
 : NUMBER OF SEQUENCES: 797  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 : STREET: 3174 PORTER DRIVE  
 : CITY: PALO ALTO  
 : STATE: CALIFORNIA  
 : COUNTRY: USA  
 : ZIP: 94304  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/60/061,998  
 : FILING DATE:  
 : CLASSIFICATION:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: CERRONE, MICHAEL C.  
 : REGISTRATION NUMBER: 39,132  
 : REFERENCE/DOCKET NUMBER: PM-0006-2P  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (650) 855-0555  
 : TELEFAX: (650) 845-4166  
 : INFORMATION FOR SEQ ID NO: 596:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1033 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: cDNA  
 : IMMEDIATE SOURCE:  
 : CLONE: SPN1c601  
 : US-60-061-998-596  
 Query Match 99.1%; Score 1981.4; DB 45; Length 10333;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 TAAATCTACGACAAATAAATCAATCACTGATGCTGAGTGGTCTGACCCGCGCTCA 60  
 DB 9718 TAAATCTACGACAAATAAATCAATCACTGATGCTGAGTGGTCTGACCCGCGCTCA 9659  
 QY 61 TGCCCAACCTAATGATATATCCACAGATTGGTTAAGCAATGTTCTTATGGAAGCA 120  
 DB 9658 TGCCCAACCTAATGATATATCCACAGATTGGTTAAGCAATGTTCTTATGGAAGCA 9599  
 QY 121 TCGCTTCTTGACCAACAGGGGATGATACCATCCGATCCTTGGAGCTTTCTTGCGCA 180  
 DB 9598 TCGCTTCTTGACCAACAGGGGATGATACCATCCGATCCTTGGAGCTTTCTTGCGCA 9539  
 QY 181 TCTGCAAGCAATTCCTCCAGAGTGATCACTCACTCAACCAAGTTGATTAGTTGAC 240  
 DB 9538 TCTGCAAGCAATTCCTCCAGAGTGATCACTCACTCAACCAAGTTGATTAGTTGAC 9479  
 QY 241 TTACTTTTCAACTGAGCTCCGACGACGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300  
 DB 9478 TTACTTTTCAACTGAGCTCCGACGACGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 9419

OY	301	AGATATTCAGTTTGAACAAAAACCAACGACGAAGAAATCTTGACCTACTATATATAA	360
Db	9418	AGCATTTCAGTTTGAACAAAAACCAACGACGAAGAAATCTTGACCTACTATATATAA	9359
OY	361	GGTTCATATGTCATTAATGGGAACATATGGAATGACAGACAGCTCAAACTACTATGTAA	420
Db	9358	GGTTCATATGTCATTAATGGGAACATATGGAATGACAGACAGCTCAAACTACTATGTAA	9299
OY	421	AGACCTCAATTAATTTAGTTTACCTCAGTTAGCCTTGTGGTGGAAATGCTCAGGCACC	480
Db	9298	AGACCTCAATTAATTTAGTTTACCTCAGTTAGCCTTGTGGTGGAAATGCTCAGGCACC	9239
OY	481	AAACCAATATGACCCCTATTACATCCAGAGACGCCCAAGACGCCGAAACTTGTCTT	540
Db	9238	AAACCAATATGACCCCTATTACATCCAGAGACGCCCAAGACGCCGAAACTTGTCTT	9179
OY	541	ATTCGAATGAAAAATCAAGGCTACATCTCTGTGAAAGCTGTGGAAGAAGCAATAC	600
Db	9178	ATTCGAATGAAAAATCAAGGCTACATCTCTGTGAAAGCTGTGGAAGAAGCAATAC	9119
OY	601	ACCAATTTACTGATGACTACAAAGTCTCAAAATCAGACAGTAATATTAACCTGTTACATGA	660
Db	9118	ACCAATTTACTGATGAGCTACAAAGTCTCAAAATCAGACAGTAATATTAACCTGTTACATGA	9059
OY	661	TAAATTCCTCAAGGAAGCTATCAATCAATCAAGTTGAGAAGAAAGGCTATTAACCTACTAC	720
Db	9058	TAAATTCCTCAAGGAAGCTATCAATCAATCAAGTTGAGAAGAAAGGCTATTAACCTACTAC	8959
OY	721	AACGGGATGGATGTCACACAAATGTACGACGAAGAGCTCAAAAACATCTGTGGGAT	780
Db	8998	AACGGGATGGATGTCACACAAATGTACGACGAAGAGCTCAAAAACATCTGTGGGAT	8939
OY	781	TTACATATACAGAGCAATACCTTGCTCATCCAGACGATGATTTGCAAGTGGTTACCAT	840
Db	8938	TTACATATACAGAGCAATACCTTGCTCATCCAGACGATGATTTGCAAGTGGTTACCAT	8879
OY	841	TGTTGATGTTTCTACGGTAAAGTCATTGCCAGCTAGGACGACCCATCACTCAATAA	900
Db	8878	TGTTGATGTTTCTACGGTAAAGTCATTGCCAGCTAGGACGACCCATCACTCAATAA	8819
OY	901	TGTTTCCTGGGAATTACCAACGACATAGAAACAAACCGCGCTGGGATCAACTATGAA	960
Db	8818	TGTTTCCTGGGAATTACCAACGACATAGAAACAAACCGCGCTGGGATCAACTATGAA	8759
OY	961	ACCGATCACAGACTACTGCTGCTGCTGGAGTACGGTGTCTACGATTAACGCTACTAT	1020
Db	8758	ACCGATCACAGACTACTGCTGCTGCTGGAGTACGGTGTCTACGAGTCAACTGCCACTAT	8659
OY	1021	CGTTACGATGAGCCCTTATACACTACCCCTGGGCAAAATACTCCTGTTTAACTGGGATG	1080
Db	8658	CGTTACGATGAGCCCTTATACACTACCCCTGGGCAAAATACTCCTGTTTAACTGGGATG	8639
OY	1081	GGGCTACTTTGGCAACATCACTTGCAATRGCCCTGCAACAATGCGGAAAGTCCAGC	1140
Db	8638	GGGCTACTTTGGCAACATCACTTGCAATRGCCCTGCAACAATGCGGAAAGTCCAGC	8579
OY	1141	CGTGGAAACCTTAACAAGGTGGGACTCAACCGCGGCAAGATTCTCTAAATGGTCTAG	1200
Db	8578	CGTGGAAACCTTAACAAGGTGGGACTCAACCGCGGCAAGATTCTCTAAATGGTCTAG	8519
OY	1201	AATGCACTACCCAGATTCTACTACTCAATATGCAATTCCTAAAGTACACAGACGATCAG	1260
Db	8518	AATGCACTACCCAGATTCTACTACTCAATATGCAATTCCTAAAGTACACAGCAGATCAG	8459
OY	1261	CAAAAAATTTGGAGCAAGTATGTAAGAAAGATGGCTGTACGCTGCTTTGCCAAATGG	1320
Db	8458	CAAAAAATTTGGAGCAAGTATGTAAGAAAGATGGCTGTACGCTGCTTTGCCAAATGG	8399
OY	1321	TGGAACCTTACTATTAACCAATGTATATCAATAAAGTCGCTTTAGTGAATGGAGTAAGA	1380
Db	8398	TGGAACCTTACTATTAACCAATGTATATCAATAAAGTCGCTTTAGTGAATGGAGTAAGA	8339
OY	1381	AGAGTTCCTTAATCTCGGAAGCTGCTGCATGAAGAAAGACAGGCTATATGATACCGA	1440

```

Db      8338      AGAGTCTCTAATGTCGGAACTCGTGCCATGAAAGAACACACAGCCTATATGATGACCA 827
      |||||||
Qy      1441      CATGATGAAAAACACTCTTGAATTAATGAACTGACGAAATGCTATCTTGCTGCC 1500
      |||||||
Db      8278      CATGATGAAAAACACTCTTGAATTAATGAACTGACGAAATGCTATCTTGCTGCC 821
      |||||||
Qy      1501      TCAGGCTGTGTAAACAGGAACCTCTTAACATAACAGACGAGGAAATTGAAACCATCA 156
      |||||||
Db      8218      TCAGGCTGTGTAAACAGGAACCTCTTAACATAACAGACGAGGAAATTGAAACCATCA 815
      |||||||
Qy      1551      GACCTCTCAATTTGTAGCACTGATGAACTATTTGGCTATACGGGTAAATATTCAT 162
      |||||||
Db      8158      GACCTCTCAATTTGTAGCACTGATGAACTATTTGGCTATACGGGTAAATATTCAT 809
      |||||||
Qy      1621      GCGGTATGAGCAGAGCTATTTCAACGCTGACACCACTTGTAGGCATGCGCTACGGT 168
      |||||||
Db      8098      GCGGTATGAGCAGAGCTATTTCAACGCTGACACCACTTGTAGGCATGCGCTACGGT 803
      |||||||
Qy      1681      CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGGAAGGAAAGCAATCAGA 174
      |||||||
Db      8038      CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGGAAGGAAAGCAATCAGA 797
      |||||||
Qy      1741      TTGGAATATACCAAGGGGCTCTACAGAATGGAATTTGTTATTTAAAAATGCTGCTG 180
      |||||||
Db      7978      TTGGAATATACCAAGGGGCTCTACAGAATGGAATTTGTTATTTAAAAATGCTGCTG 791
      |||||||
Qy      1801      TTTCACGCGGAACGACCTGCTCCACAAACACCCCATCAAGTGAAGTTCACAGCTCATC 186
      |||||||
Db      7918      TTTCACGCGGAGCTCACCCTGCTCCACAAACACCCCATCAAGTGAAGTTCACAGCTCATC 785
      |||||||
Qy      1861      ATCAGATAGTTCAACTTCACAGCTCAGCTCAACCACTCCAGACCAAAATATAGTAGAC 192
      |||||||
Db      7858      ATCAGATAGTTCAACTTCACAGCTCAGCTCAACCACTCCAGACCAAAATATAGTAGAC 779
      |||||||
Qy      1921      TACCAATCCTTAACAATAATAGCGAAACATCAATACAAACCCCTGATATCAAAAAATGAA 198
      |||||||
Db      7798      TACCAATCCTTAACAATAATAGCGAAACATCAATACAAACCCCTGATATCAAAAAATGAA 773
      |||||||
Qy      1981      TCGTCACACGACACACCA 1999
      |||||||
Db      7738      TCGTCACACGACACACCA 7720
      |||||||

RESULT      12
US-60-068-175-596/C
: Sequence 596, Application US/60068175
: GENERAL INFORMATION:
: APPLICANT: Lagace, Robert E.
: APPLICANT: Corley, Neil C.
: APPLICANT: Russo, Frank D.
: APPLICANT: Hann, Amy L.
: APPLICANT: Heath, Joe D.
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
: NUMBER OF SEQUENCES: 1175
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 PORTER DRIVE
: CITY: PALO ALTO
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94304
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/60/068,175
: FILING DATE: HEREWITH
: CLASSIFICATION:
:

```





```

; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
; NUMBER OF SEQUENCES: 713
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/14436
; FILING DATE: 15-AUG-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,022
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmil, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50533
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
;
; TELEX:
; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; PCT-US97-14436-139

```

Query Match 40.58; Score 808.6; DB 1; Length 2172;  
Best Local Similarity 99.48; Pred. No. 5.1e-218;  
Matches 822; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```

Oy 1174 CGCCAAGACTTCTCTAATAGTGTAGAGATCGACTACCAAGTATTCACACTCAATGCG 1233
    |||||||
Db 1 CGCCAAGACTTCTCTAATAGTGTAGAGATCGACTACCAAGTATTCACACTCAATGCG 60
Oy 1234 CATTTCAGTAAACACAAACGCAATCAGCAAAAAAATATGAGCAAGTGAAGATGGC 1293
    |||||||
Db 61 CATTTCAGTAAACACAAACGCAATCAGCAAAAAAATATGAGCAAGTGAAGATGGC 120
Oy 1294 TGGTGGCTTAGCGTCCCTTTCGAATGTGTGGAAGTCTAATGTGGAAGTGTGCGCATGAA 1353
    |||||||
Db 121 TGGTGGCTTAGCGTCCCTTTCGAATGTGTGGAAGTCTAATGTGGAAGTGTGCGCATGAA 180
Oy 1354 AGTCGCTTAGTGTGAGTGGAGTGAAGAAAGATTCTAATGTGGAAGTGTGCGCATGAA 1413
    |||||||
Db 181 AGTCGCTTAGTGTGAGTGGAGTGAAGAAAGATTCTAATGTGGAAGTGTGCGCATGAA 240
Oy 1414 GGAAGACAGACGCTATATGTATGATGACCGACATGATGAAAAACAGTCTTGAAGTGAAGTGG 1473
    |||||||
Db 241 GGAAGACAGACGCTATATGTATGATGACCGACATGATGAAAAACAGTCTTGAAGTGAAGTGG 300
Oy 1474 ACG-AAATGGCTATCTGCTGGTGGCTCCCTCAGAGCTGTGTAAGACGAGAACTCTAAGTATA 1592
    |||
Db 301 ACGAAATGGCTATCTGCTGGTGGCTCCCTCAGAGCTGTGTAAGACGAGAACTCTAAGTATA 360
Oy 1533 CAGACGAGAAATGTAAGAACACATCAAGACCTCTCAATTGTGTGACCTGATGAAGTAT 1592
    |||||||
Db 361 CAGACGAGAAATGTAAGAACACATCAAGACCTCTCAATTGTGTGACCTGATGAAGTAT 420

```

```

Oy 1593 TTGCTGAGCTATACGCGTAATATTCATATGCGTGTATGAGACAGGCTATTTACCGTCTGA 1652
    |||||||
Db 421 TTGCTGAGCTATACGCGTAATATTCATATGCGTGTATGAGACAGGCTATTTACCGTCTGA 480
Oy 1653 CACACCTGTAGAGCAATGCGCTTACGGCTGTGCAAGTTTACCGTGTATGATGACT 1712
    |||||||
Db 481 CACACCTGTAGAGCAATGCGCTTACGGCTGTGCAAGTTTACCGTGTATGATGACT 540
Oy 1713 ACCTGTGTAAGAGCAATCCAGAAATGTAATATACAGAGGGGCTGTACAGAAATG 1772
    |||||||
Db 541 ACCTGTGTAAGAGCAATCCAGAAATGTAATATACAGAGGGGCTGTACAGAAATG 600
Oy 1773 GAGAAATGCTATTTAAAAATGCTGCTGCTTCACGTGGAATCACTGCTCCACAAAC 1832
    |||
Db 601 GAGAAATGCTATTTAAAAATGCTGCTGCTTCACGTGGAATCACTGCTCCACAAAC 660
Oy 1833 CCCCATCACTGAAAGTTCAGATCATCATCATAGTATGATTCACATTCACAGTGTGCTCAA 1892
    |||||||
Db 661 CCCCATCACTGAAAGTTCAGATCATCATCATAGTATGATTCACATTCACAGTGTGCTCAA 720
Oy 1893 CCACTCCAGCACAATAATATAGTACGACTACCAATCCTTAACAATAATACCAACATCAA 1952
    |||||||
Db 721 CCACTCCAGCACAATAATATAGTACGACTACCAATCCTTAACAATAATACCAACATCAA 780
Oy 1953 ATACACCCCTGATCAACAATAATCAGAAATCCTCAACGACACAAACCA 1999
    |||||||
Db 781 ATACACCCCTGATCAACAATAATCAGAAATCCTCAACGACACAAACCA 827

```

RESULT 15  
US-08-911-503-139  
; Sequence 139, Application US/08911503  
; GENERAL INFORMATION:  
; APPLICANT: Black, Michael  
; APPLICANT: Hodgson, John  
; APPLICANT: Knowles, David  
; APPLICANT: Lonetto, Michael  
; APPLICANT: Nicholas, Richard  
; APPLICANT: Stodola, Robert  
; TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,  
; NUMBER OF SEQUENCES: 713  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/911,503  
; FILING DATE: 15-AUG-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/024,022  
; FILING DATE: 16-AUG-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gimmil, Edward R  
; REGISTRATION NUMBER: 38,891  
; REFERENCE/DOCKET NUMBER: P50533  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-4478  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 139:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2172 base pairs  
; TYPE: nucleic acid



STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-08-911-503-139

Query Match 40.5%; Score 808.6; DB 13; Length 2172;

Best Local Similarity 99.4%; Pred. No. 5,1e-218;  
Matches 822; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1174 CGCCAGAGCTTTCCTTAATGCTAGGAATGAGATACCAAGTAATGCTACTACTCAAAATGC 1233  
Db 1 CGCCAGAGCTTTCCTTAATGCTAGGAATGAGATACCAAGTAATGCTACTACTCAAAATGC 60  
QY 1234 CATTTCAGTAACACAAACCAATCAGACAAAATAATGAGCAAGTAGTGAAGAAAGATGCC 1293  
Db 61 CATTTCAGTAACACAAACCAATCAGACAAAATAATGAGCAAGTAGTGAAGAAAGATGCC 120  
QY 1294 TGGTGGTTAGCGTCCCTTTGCAATGGTGGAACTTACTATAAACCAATGTATATCCATA 1353  
Db 121 TGCTGCTTACGCTGCCCTTTGCAATGGTGGAACTTACTATAAACCAATGTATATCCATA 180  
QY 1354 AGTCGCTTTAGTATGAGAGAGTGAAGAAAGATTCTTAATGTCGGAAGTGTGCCATGAA 1413  
Db 181 AGTCGCTTTAGTATGAGAGAGTGAAGAAAGATTCTTAATGTCGGAAGTGTGCCATGAA 240  
QY 1414 GGAACGACAGCCTATATGATGACCGACATGATGAAAAACAGTCTTGACTTATGGAAGTGG 1473  
Db 241 GGAACGACAGCCTATATGATGACCGACATGATGAAAAACAGTCTTGACTTATGGAAGTGG 300  
QY 1474 ACG-AAATGCTATCTGCTGGCTCCGCTAGGCTGTAAACGGAACCTCTAATATA 1532  
Db 301 ACGAAATGCTATCTGCTGGCTCCGCTAGGCTGTAAACGGAACCTCTAATATA 360  
QY 1533 CAGACGAGAAATTTGAAGAACACATCAAGACCTCTCAATTTGTAGACACTGATGAATAT 1592  
Db 361 CAGACGAGAAATTTGAAGAACACATCAAGACCTCTCAATTTGTAGACACTGATGAATAT 420  
QY 1593 TTGCTGGCTATACGCGTAAATATTCATATGCTGATGAGACAGGCTATTTAACCCTGTA 1652  
Db 421 TTGCTGGCTATACGCGTAAATATTCATATGCTGATGAGACAGGCTATTTAACCCTGTA 480  
QY 1653 CACCACCTTGAAGCAATGGCCTTACGGCTGCGCAAGTTTACCGCTATATGATGACT 1712  
Db 481 CACCACCTTGAAGCAATGGCCTTACGGCTGCGCAAGTTTACCGCTATATGATGACT 540  
QY 1713 ACCTGCTGAAGAGCAATCCAGAAATTTGAATATACAGAGGGGCTCTACAGAATG 1772  
Db 541 ACCTGCTGAAGAGCAATCCAGAAATTTGAATATACAGAGGGGCTCTACAGAATG 600  
QY 1773 GAGAAATTCGATTTAAAAATGGTCTGTTCTACGTGGAACCTCACTGCTCCACACAAC 1832  
Db 601 GAGAAATTCGATTTAAAAATGGTCTGTTCTACGTGGAACCTCACTGCTCCACACAAC 660  
QY 1833 CCCCATCAACTGAAGTTCAAGCTCATCATCAGATTAATTCAACTTCAAGCTTACAGTCAA 1892  
Db 661 CCCCATCAACTGAAGTTCAAGCTCATCATCAGATTAATTCAACTTCAAGCTTACAGTCAA 720  
QY 1893 CCACTCCACACCAAAATATAGTACGACTCAATCTTAACAAATATATGCAACATCAA 1952  
Db 721 CCACTCCACACCAAAATATAGTACGACTCAATCTTAACAAATATATGCAACATCAA 780  
QY 1953 ATACAAACCCCTGATCAACAAATCAGAAATCCTCAACAGCACAACCA 1999  
Db 781 ATACAAACCCCTGATCAACAAATCAGAAATCCTCAACAGCACAACCA 827

Search completed: June 13, 2002, 14:43:46  
Job time: 7603 sec

**THIS PAGE BLANK (USPTO)**



Db 1 TAAATCTAGCAATAAAAATCAATCTAGTGTGGTCTGAAAGCCGCGTCAA 60  
Qy 61 TCCCAAGCTAATGATATCCACAGATTTGGTAAGGCATGTTTCTATCGAAGACA 120  
Db 61 TCCCAAGCTAATGATATCCACAGATTTGGTAAGGCATGTTTCTATCGAAGACA 120  
Qy 121 TCGCTTCTGACACACAGGGGGGATGATACATCCGATCTGAGAGCTTTCTGGCAA 180  
Db 121 TCGCTTCTGACACACAGGGGGGATGATACATCCGATCTGAGAGCTTTCTGGCAA 180  
Qy 181 TCTGAAAGCAATTCCTCCCAAGGTGATGATCACTCCACCCCAAGCTGATTAAGTGC 240  
Db 181 TCTGAAAGCAATTCCTCCCAAGGTGATGATCACTCCACCCCAAGCTGATTAAGTGC 240  
Qy 241 TTAATTTCAACTTGCAGCTTCGACAGCACTATTTCTGTAAGGCTCAGAGAGCTTGGT 300  
Db 241 TTAATTTCAACTTGCAGCTTCGACAGCACTATTTCTGTAAGGCTCAGAGAGCTTGGT 300  
Qy 301 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATTAATTA 360  
Db 301 AGCGATTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTATATTAATTA 360  
Qy 361 GGTCTACATGTGTATGGGAATGATGAATGAGACAGAGCTCAAACTACTATGTGA 420  
Db 361 GGTCTACATGTGTATGGGAATGATGAATGAGACAGAGCTCAAACTACTATGTGA 420  
Qy 421 AGACCTCATATTAATTAAGTTTACCTAGTTAGCTTGGCTGGGAATGCTCAGAGACC 480  
Db 421 AGACCTCATATTAATTAAGTTTACCTAGTTAGCTTGGCTGGGAATGCTCAGAGACC 480  
Qy 481 AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCCAAGACCGGAACCTTGGCT 540  
Db 481 AAACCAATATGACCCCTATTCACATCCAGAAAGCAGCCCAAGACCGGAACCTTGGCT 540  
Qy 541 ATCTGAATGAAATTCAGAGCTACATCTCTGTGAACAGTGTGAAGAAAGCTCAATAC 600  
Db 541 ATCTGAATGAAATTCAGAGCTACATCTCTGTGAACAGTGTGAAGAAAGCTCAATAC 600  
Qy 601 ACCAATTCAGTATGATGATCAAAAGTCTCAAAATGATGATTAATTCCTGCTTACATGA 660  
Db 601 ACCAATTCAGTATGATGATCAAAAGTCTCAAAATGATGATTAATTCCTGCTTACATGA 660  
Qy 661 TAATTAACCTCAAGAGATCATCAATCAAGTGAAGAAAGAGCTCAAACTACTACTAC 720  
Db 661 TAATTAACCTCAAGAGATCATCAATCAAGTGAAGAAAGAGCTCAAACTACTACTAC 720  
Qy 721 AACTGGATGATGATCTCTACAAATGATGATGATGATGATGATGATGATGATGAT 780  
Db 721 AACTGGATGATGATCTCTACAAATGATGATGATGATGATGATGATGATGATGAT 780  
Qy 781 TTACATATCAGACGATAGTGTGCTATCCAGAGATGATGATGATGATGATGATGAT 840  
Db 781 TTACATATCAGACGATAGTGTGCTATCCAGAGATGATGATGATGATGATGATGAT 840  
Qy 841 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Db 841 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Qy 901 TGTTCCTTGGGAATTAACGAGAGATGATGATGATGATGATGATGATGATGATGAT 960  
Db 901 TGTTCCTTGGGAATTAACGAGAGATGATGATGATGATGATGATGATGATGATGAT 960  
Qy 961 ACCGATTCAGACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
Db 961 ACCGATTCAGACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
Qy 1021 CGTTACAGATGAGCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080  
Db 1021 CGTTACAGATGAGCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080  
Qy 1081 GGGCTACTTGGCAACATCAGCTTGCATATAGCCCTGCAACAAATGCGAAAGCTGCCAGC 1140  
Db 1081 GGGCTACTTGGCAACATCAGCTTGCATATAGCCCTGCAACAAATGCGAAAGCTGCCAGC 1140

Qy 1141 CGTGAAGCTCTAAACAAAGTGGGACTCAACCGGCCCAAGACTTTCTAAATGCTTAGG 1200  
Db 1141 CGTGAAGCTCTAAACAAAGTGGGACTCAACCGGCCCAAGACTTTCTAAATGCTTAGG 1200  
Qy 1201 AATGCACTACCCAGATTTACTACTCAAAATGCCATTTCAAGTAAACACACCAATCAGA 1260  
Db 1201 AATGCACTACCCAGATTTACTACTCAAAATGCCATTTCAAGTAAACACACCAATCAGA 1260  
Qy 1261 CAAAAAATATGAGCAAGTAGGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
Db 1261 CAAAAAATATGAGCAAGTAGGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
Qy 1321 TGGAACTTACTATTAACCAATGATATATCCATTAATGCTTAACTGATGAGAGTGA 1380  
Db 1321 TGGAACTTACTATTAACCAATGATATATCCATTAATGCTTAACTGATGAGAGTGA 1380  
Qy 1381 AGAGTTCTTAATGTCGGAACCTGTCATGAAAGCAAGAGAGAGAGAGAGAGAGAGAG 1440  
Db 1381 AGAGTTCTTAATGTCGGAACCTGTCATGAAAGCAAGAGAGAGAGAGAGAGAGAGAG 1440  
Qy 1441 CATGATGAAACAGCTTACTATGATGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 1500  
Db 1441 CATGATGAAACAGCTTACTATGATGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 1500  
Qy 1501 TCAGAGCTGTTAAACAGAGACCTTAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
Db 1501 TCAGAGCTGTTAAACAGAGACCTTAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
Qy 1561 GACCTCTCAATTTGATGACAGCTGATGAATTTGCTGCTATATACGCTTAATTTCAAT 1620  
Db 1561 GACCTCTCAATTTGATGACAGCTGATGAATTTGCTGCTATATACGCTTAATTTCAAT 1620  
Qy 1621 GGGCTATGAGAGAGCTTATTAACCGTCTGACACACCTGTAAGCAATGAGCTTAGGCT 1680  
Db 1621 GGGCTATGAGAGAGCTTATTAACCGTCTGACACACCTGTAAGCAATGAGCTTAGGCT 1680  
Qy 1681 CGCTGCCAAAGTTTACCGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
Db 1681 CGCTGCCAAAGTTTACCGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
Qy 1741 TTGGAATATACAGAGGAGCTCTACAGAAATGAGAAATTCGATTTAAATGAGTCTG 1800  
Db 1741 TTGGAATATACAGAGGAGCTCTACAGAAATGAGAAATTCGATTTAAATGAGTCTG 1800  
Qy 1801 TTCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
Db 1801 TTCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
Qy 1861 ATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
Db 1861 ATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
Qy 1921 TACCAATCTTAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAAT 1980  
Db 1921 TACCAATCTTAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAAT 1980  
Qy 1981 TCCCTAACACAGACACA 1999  
Db 1981 TCCCTAACACAGACACA 1999

RESULT 2  
US-10-121-120-27/c  
; Sequence 27, Application US/10121120  
; GENERAL INFORMATION:  
; APPLICANT: Bergeron, Michel G.  
; APPLICANT: Ouellette, Marc  
; APPLICANT: Roy, Paul H.  
; TITLE OF INVENTION: Specific and Universal Probes and Amplification  
; TITLE OF INVENTION: Primers  
; TITLE OF INVENTION: To Rapidly Detect and Identify Common Bacterial  
; TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical

```

: TITLE OR INVENTION:  Specimens for Routine Diagnosis in Micro
:
: FILE REFERENCE:  12287.31
:
: CURRENT APPLICATION NUMBER:  US/10/121,120
:
: CURRENT FILING DATE:  2002-04-11
:
: PRIOR APPLICATION NUMBER:  09/4452,599
:
: PRIOR FILING DATE:  1999-12-01
:
: PRIOR APPLICATION NUMBER:  08/304,732
:
: PRIOR FILING DATE:  1994-09-12
:
: NUMBER OF SEQ ID NOS:  177
:
: SOFTWARE:  Patentin Ver. 2.1
:
: SEQ ID NO 27
:
: LENGTH: 9100
:
: TYPE:  DNA
:
: ORGANISM:  Haemophilus influenzae
:
: OS-10-121-120-27

```

Query March	5.9%	Score 117.2;	DB 6;	Length 9100;
Best Local Similarity	50.3%	Pred. No. 1.5e-22;		
Matches 349;	Conservative 0;	Mismatches 333;	Indels 12;	Gaps 2

QY	22	TCACATCTATGCTGACATTGGGTTCTCAAGCCGCGCAATGCCAAGCTAATGATATTC	81
Db	3498	TAAATTATTTGGCGGAAGTGGGTGAGCAACGCCGTATTCAGTAAATTAAGCCGATGTGC	34299
QY	82	CACAGATTGGTTTAAGCAATCGTTTCTATGCAAGACCATGCGTTCTTGACCAAGGG	141
Db	3438	ACAACGGTTTAAATGACGATCTTTTATACGCGAGGAAAGCATGCTGTTTACGATCATCACGG	3379
QY	142	GATTGATACCATCCGATATCCTGGGACCTTCTTGGGCAATCGCAAA---GCAATTCCT	198
Db	3378	AATAGACCTTATGGCATTTGCCATTTGGTGTGCGAGTAAATGCGGTGCATC	3319
QY	199	CCAAGTGTATCAACTCTGACCCACAGTTGATTAAGTTGACTTACTTTGCACTTCAC	258-
Db	3318	ACAAGGCGCAATGACATTACTCAACAATTAAGCGGTAACTTTTCTTA-----AC	3288
QY	259	TTCCGACACGACTATTTTCTGTAAAGCTCAGAAAGCTGGTGGTACGATTACAGACA	318
Db	3267	CTAGAAAAAACCATTTATTCGTAAACCTCTGTAAACCGGTGCGGTGAGAAATGAA	3208
QY	319	AAAAGCACCAACCAAGAATCTTGACTGACTATATTAATAAGTCTCAATGCTAATAG	378
Db	3207	TACTCTCAACAACCAAGAAATATTACAGCTTATTTAAACAATCTTTTAAAGGTATCG	3148
QY	379	GAACATATGATCGACAGACAGCTCAAAATCTATGTGTAAAGCTCAATTAATTAA	438
Db	3147	TTCTTATGATGGTTGACGCGGACAGCAACAACTATTCGGTAAATCATTTGAATGATTC	3088
QY	439	TTTACCTCATTTAGCCCTGCTGGCTGGAATGGCTGACGACCACCAACCAATATGACCCCTA	498
Db	3087	CTTATCGGAATGGCATTTATTTCTGTGTTTACTTAAGCACCTTCAACAATGAACCCGCT	3028
QY	499	TTGCATATCGAAGACCCCAAGACCGCGAAACTGGTCTTATCTGAAATGAAAAATCA	558
Db	3027	TTTATCTTTAAACGTTTCAGAAAGAACGCCGCAATGAGTGCTTAACCGTATGTAATGA	2968
QY	559	AGGCTATATCTGCTGGAACAGTATGAAAGAACGACTCAATACCAATTTACTGATGACT	618
Db	2967	AAATATCATCGCAAGAAAGAAATATGATCTCTCATTTGAAAGAACCGCATTTGTGGGAGCTA	2908
QY	619	ACAAGCTCAATTCAGCAAGTAAATTAACCTGCTTACATATGTAATTAATCTCAAGCAAGT	678
Db	2907	TCAAGGCGCAAAATTTGAATTTTCGACCGCATTTATGCACTGAAATATGATGCTCAAGAAAT	2848
QY	679	CATCATCAAGTTGAAGAAAGAAACAGGCTATTAAC	712
Db	2847	GGTGGCTGTTTGGGCAAGAAATCTTTACCC	2814

RESULT 3  
US-09-540-209B-3719  
; Sequence 3719, Application US/09540209B

```

: GENERAL INFORMATION:
: APPLICANT: Gairy L. Breton
: TITLE OF INVENTION: NOCLETIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
: TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 2709.1001-001
: CURRENT APPLICATION NUMBER: US/09/540,209B
: CURRENT FILING DATE: 2000-04-04
: NUMBER OF SEQ ID NOS: 10444
: SEQ ID NO 3719
: LENGTH: 2370
: TYPE: DNA
: ORGANISM: B.fragilis
: US-09-540-209B-3719

```

Query Match	3.0%	Score	59.2	DB	5	Length	2370
Best Local Similarity	46.9%	Pred	No.2.2e-06				
Matches	261	Conservative	0	Mismatches	283	Indels	12
						Gaps	2

QY	42	GTTCCTGAACGGCCGGCTGAATGCCCAAGCCTATATATATTTCCACAGATTTGGTTGAAGCAA	101
Db	233	glaagagaaatcgtgtatacaaccttataatgaactttcaaccaacatctgtcaatgact	292
QY	102	TGCTTTCTATCGAAGAACCATTCGCTTTCTTGACCAACAAGGGGATTTGATACCATCCGTATCC	161
Db	293	tgtatcgcaacgagagacgttccttcgaacaacatctcgtgtatcgtgtcgcgaagcgtga	352
QY	162	TGGGAGCTTTCCT-----GCGCAATCTGCAAAAGCAATTCCTCCCAAGGTGGATCAACTC	215
Db	353	taagtgtcgtgtgttaagcgttgcgtatctgcgtatgcagaanaatgcaggttgaggcagatcac	412
QY	216	TGACCCCAACAGTTGATTA-----GTGGACTTACTTTTCAACTTCGTGACCTCGACACGA	269
Db	413	tttcaacaacagctgcgaagaacatctgttttcgcgaagaagttgcgaagaataagctgcgagc	472
QY	270	CTATTTCTCGTAAGGCTCAGAGAACTTGTTAGTCGATTCAGTTAGAACAAAAAGCAACA	329
Db	473	gcctgtttcagaagccgatagagtggtgtgttcgcgttaaacctgtaacgtttattacaa	532
QY	330	AGCAAGAATCTTACTACTATATATAATTAAGTCTACATGTTAATGGGAACATAAGAA	389
Db	533	aggaagaacatttbgatatactatccataaaattgacttcctgtaataatgcagtagaa	592
QY	390	TGCGAGACGACGCTCAAAAGCTACTATATGTTAAAGCCCAATATAATTAAGTTTAACTCCAGT	449
Db	593	ttaaacgcgtctcataaccatttgcgtatgcgaaccacaagaactcgtgaataatagaacag	652
QY	450	TAGCCTTGCCTGGCGTGAATCCTCTCAGGACCAACAACATATGACCCCTATTGCATCATCAG	509
Db	653	ctgttaacgctgcgtatcgtgtatgtgtaaaaatcccttcgctttacaatccggtgcgtccaag	712
QY	510	AAGCAGCCCAAGACCGCCGAACCTTGTTCTTATCTGAATGAAAAATCAAGCTCATCTCT	569
Db	713	agcgctcgcgcgacgcagcaatatacagtgctagatcaaatlgaaagaagcccgatatatca	772
QY	570	CTGCTGAACAGTATGA	585
Db	773	caagcgaagaagatgta	788

```

RESULT      4
US-09-673-476-206
: Sequence 206, Application US/09673476
: GENERAL INFORMATION:
: APPLICANT: COLE, STEWART
: APPLICANT: BUCHRIEISER-BROSCH, ROLAND
: APPLICANT: GORDON, STEPHEN
: APPLICANT: BILLALUT, ALAIN
: TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
: TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
: TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
: TITLE OF INVENTION: MYCOBACTERIA.
: FILE REFERENCE: 05394.0011-00000

```

```

: CURRENT APPLICATION NUMBER: US/09/673,476
: CURRENT FILING DATE: 2002-03-29
: PRIOR APPLICATION NUMBER: PCT/IB99/00740
: PRIOR FILING DATE: 1999-04-16
: PRIOR APPLICATION NUMBER: 09/060,756
: PRIOR FILING DATE: 1998-04-16
: NUMBER OF SEQ ID NOS: 743
: SOFTWARE: PatentIn Ver. 2.2
: SEQ ID NO 206
: LENGTH: 428
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (280)
: OTHER INFORMATION: a, t, c or g
US-09-673-476-206
```

```

Query Match          2.8%; Score 55.2; DB 5; Length 428;
Best Local Similarity 50.8%; Pred. No. 1.5e-05;
Matches 132; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
```

```

QY 281 AAGGCTCAGGAAGCTTGGTTAGCGATTGATAGCAAAAAAGCAAGCAAGAATC 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7 aagctcgagatccgagatgcgactcgcgcgcaagacccctcacaatctgaatc 66
QY 341 TTGACCTACTATATTAATAGGTCTACATGCTATATGGAACTATGAACTGACAGCA 400
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 67 ctgaccgacatctgacacctgctcgtcgcgaataactcgttcgcaagagcg 126
QY 401 GCTCAAACTACTATGTAAGAGCCTCAATATTAAGTTTACCTAGCTTGGCTG 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 127 ggcgaagaagctcttgcgcatacaagcgctcgaacctgaatgcaagaagcgctgctg 186
QY 461 GCTGAATGCTCTCAGGACCAACCAATATGACCCCTATTCATTCAGAACGACCCCA 520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 187 gccgcatactgtaacatcgaccagcagcgtcacaaccgtaacaccccgagcgcgctg 246
QY 521 GACCCCGCAACTTGGCTCT 540
    ||| ||| ||| ||| |||
DB 247 gcccgcggaacgtgctct 266
```

```

RESULT 5
US-09-673-476-626
: Sequence 626, Application US/09673476
: GENERAL INFORMATION:
: APPLICANT: COLE, STEWART
: APPLICANT: BUCHRIEISER-BROSCH, ROLAND
: APPLICANT: GORDON, STEPHEN
: APPLICANT: BILLAUFT, ALAIN
: TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
: TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
: TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
: FILE REFERENCE: 05394.0011-00000
: CURRENT APPLICATION NUMBER: US/09/673,476
: CURRENT FILING DATE: 2002-03-29
: PRIOR APPLICATION NUMBER: PCT/IB99/00740
: PRIOR FILING DATE: 1999-04-16
: PRIOR APPLICATION NUMBER: 09/060,756
: PRIOR FILING DATE: 1998-04-16
: NUMBER OF SEQ ID NOS: 743
: SOFTWARE: PatentIn Ver. 2.2
: SEQ ID NO 626
: LENGTH: 363
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
US-09-673-476-626
```

```
Query Match          2.6%; Score 51.8; DB 5; Length 363;
```

```

Best Local Similarity 52.6%; Pred. No. 0.00013;
Matches 113; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 326 ACCAAGCAAGAAATCTTACCTACTATATTAATAGTCTACATGTCTAATGGCAAGAT 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 26 acaaatctgaaatcttgcacccgatacttgaacctgtctcgttgcgaataactcgtc 85
QY 386 GGAATGCAGACAGCAGCTCAAAATCTACTATGTTAAAGCCTCAATATTAATTTAGTTACT 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86 ggcgtcggaagcgcgcgcgaacacgtaacttgcgcatacaacgctccgactgaattgca 145
QY 446 GATTAGCCTTGCTGCTGGAATGCTTCAGGCACCAACCAATATGACCCCTATTTCAT 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 146 caagcgcgctcgtcgcgcgacatggtgcaatcgaccagcagcctcaaccgtaacccaac 205
QY 506 CCAGAAGCAGCCCAAGACCCGCGCAACTTGGCTCT 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 206 cccgacgcgcgcgtgcgcgcgcgaacgltgctct 240
```

```

RESULT 6
US-09-673-476-438
: Sequence 438, Application US/09673476
: GENERAL INFORMATION:
: APPLICANT: COLE, STEWART
: APPLICANT: BUCHRIEISER-BROSCH, ROLAND
: APPLICANT: GORDON, STEPHEN
: APPLICANT: BILLAUFT, ALAIN
: TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
: TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
: TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
: FILE REFERENCE: 05394.0011-00000
: CURRENT APPLICATION NUMBER: US/09/673,476
: CURRENT FILING DATE: 2002-03-29
: PRIOR APPLICATION NUMBER: PCT/IB99/00740
: PRIOR FILING DATE: 1999-04-16
: PRIOR APPLICATION NUMBER: 09/060,756
: PRIOR FILING DATE: 1998-04-16
: NUMBER OF SEQ ID NOS: 743
: SOFTWARE: PatentIn Ver. 2.2
: SEQ ID NO 438
: LENGTH: 223
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (136)
: OTHER INFORMATION: a, t, c or g
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (158)
: OTHER INFORMATION: a, t, c or g
: NAME/KEY: modified_base
: LOCATION: (203)
: OTHER INFORMATION: a, t, c or g
US-09-673-476-438
```

```

Query Match          2.0%; Score 40.6; DB 5; Length 223;
Best Local Similarity 48.6%; Pred. No. 0.16;
Matches 106; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
QY 281 AAGGCTCAGGAAGCTTGGTTAGCGATTGATAGCAAAAAAGCAAGCAAGAATC 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6 aagctcgagatccgagatgcgactcgcgcgaagacccctcacaatctgaatc 65
QY 341 TTGACCTACTATATTAATAGGTCTACATGCTCTAATGGAACTATGAACTGACAGCA 400
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 66 ctgaccgatactgacacctgctcgttcgcgaataactcgttcgcaagagcg 125
QY 401 GCTCAAACTACTATGTAAGAGCCTCAATATTAATTAAGTTTACCTAGCTTGGCTG 460
```







```
; TITLE OF INVENTION: MYCOBACTERIA.
; FILE REFERENCE: 05394.0011-00000
; CURRENT APPLICATION NUMBER: US/09/673,476
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/IB99/00740
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 09/060,756
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 671
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (147)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (247)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (380)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (457)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (468)
; OTHER INFORMATION: a, t, c or g
; US-09-673-476-671
```

Query Match 1.8%; Score 36.4; DB 5; Length 473;  
Best Local Similarity 51.6%; Pred. No. 3.2;

Matches 82; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```
OY 5 ATCTACGACAAATAAATCACTGATTCCTGCTGAGCGCGGCTCAATGCC 64
   || |||| || || || || || || || || || || || || || ||
DB 183 ATGTGCGACGCGAAGGACACACGATCGGTGCTGACTCCAGCGGTTCAAGTG 124
OY 65 CAAGCTAATGATATTCACACAGATTGGTTAAGCAATCGTTTCTATCGAAGCATTGCC 124
   || || || || || || || || || || || || || || || ||
DB 123 CCCTCGACAAAGATGCGCAACACGATGAAGCTGGCGATCGTCTGATTGAAGTAAGCGG 64
OY 125 TTCTTCGACCAAGGGGATTCATACCATCCGTATCCTG 163
   || || || || || || || || || || || || || || || ||
DB 63 TTGCGCGACACAGCGGCTGACTGGAAGGCGACCCCTG 25
```

```
RESULT 14
US-10-027-632-148124/c
; Sequence 148124, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
```

```
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148124
; LENGTH: 689
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-148124
```

Query Match 1.8%; Score 36; DB 6; Length 689;  
Best Local Similarity 51.2%; Pred. No. 4.8;

Matches 84; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
OY 1832 CCCCCATCACTGAAAGTCGACGTCATCATGATAGTTCACACTTCACAGTCACTAGCTCA 1891
   || || || || || || || || || || || || || || || ||
DB 360 CCCCCCTCCAGAGATATATCTCCCTCTAGAAAGATATATCTCTCTCCAGAGATTAAT 301
OY 1892 ACCACTCCAGGACCAATATAGTAGACTACCAATCTCTAACAATAATAGCAAGATCA 1951
   || || || || || || || || || || || || || || || ||
DB 300 ACTCCTCCTCCAGAGATATATCTCTCTCCAGAAAGATATATATCTCTCTCCAGAA 241
OY 1952 AATFACACCCCTGATCAACAATAATGAAATCTCTCAACACAGCACA 1995
   || || || || || || || || || || || || || || || ||
DB 240 GATAATATCTCTCTCTAGAAAGATATATCTCCCTCCAGAGAA 197
```

RESULT 15

US-09-673-476-99/c  
; Sequence 99, Application US/09673476

; GENERAL INFORMATION:

; APPLICANT: COLE, STEWART

; APPLICANT: BUCHRIESEN-BROSCH, ROLAND

; APPLICANT: GORDON, STEPHEN

; APPLICANT: BILLAUD, ALAIN

; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST

; TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED

; TITLE OF INVENTION: DNA LIBRARY, APPLICATION TO THE DETECTION OF

; TITLE OF INVENTION: MYCOBACTERIA.

; FILE REFERENCE: 05394.0011-00000

; CURRENT APPLICATION NUMBER: US/09/673,476

; CURRENT FILING DATE: 2002-03-29

; PRIOR APPLICATION NUMBER: PCT/IB99/00740

; PRIOR FILING DATE: 1999-04-16

; PRIOR APPLICATION NUMBER: 09/060,756

; PRIOR FILING DATE: 1998-04-16

; NUMBER OF SEQ ID NOS: 743

; SOFTWARE: PatentIn Ver. 2.2

; SEQ ID NO 99

; LENGTH: 348

; TYPE: DNA

; ORGANISM: Mycobacterium tuberculosis

; US-09-673-476-99

Query Match 1.8%; Score 35.8; DB 5; Length 348;  
Best Local Similarity 51.6%; Pred. No. 4.3;

Matches 82; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```
OY 5 ATCTACGACAAATAAATCACTGATTCCTGCTGAGCGCGGCTCAATGCC 64
   || |||| || || || || || || || || || || || || || ||
DB 321 ATGTGCGACGCGAAGGACACACGATCGGTGCTGACTCCAGCGGTTCAAGTG 262
OY 65 CAAGCTAATGATATTCACACAGATTGGTTAAGCAATCGTTTCTATCGAAGCATTGCC 124
   || || || || || || || || || || || || || || || ||
DB 261 CCCTCGACAAAGATGCGCAACACGATGAAGCTGGCGATCGTCTGATTGAAGTAAGCGG 202
OY 125 TTCTTCGACCAAGGGGATTCATACCATCCGTATCCTG 163
   || || || || || || || || || || || || || || || ||
```

Db 201 TTCGCCGACACAGCGGCGTGACTGGAAGGGCAACCTG 163

Search completed: June 13, 2002, 14:47:52  
Job time: 7519 sec

...

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 03:12:54 ; Search time 61.45 Seconds  
(without alignments)  
1203.827 Million cell updates/sec

Title: US-08-961-083-2  
Perfect score: 3484  
Sequence: 1 KIVDNKNQLIADLSESRVN.....TQSSNTTPOQNONPQAPQ 666

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802:\*

- 1: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*
- 2: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*
- 3: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*
- 4: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*
- 5: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*
- 6: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*
- 7: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*
- 8: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*
- 9: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*
- 10: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*
- 11: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*
- 12: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*
- 13: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*
- 14: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*
- 15: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*
- 16: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*
- 17: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*
- 18: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*
- 19: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*
- 20: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*
- 21: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*
- 22: /SIDS5/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	666	19	AAW5063
2	3466	99.5	682	17	AAW04359
3	3466	99.5	719	22	AAU37830
4	2208	63.4	420	20	AAV56106
5	1751	50.3	778	22	AAU35135
6	1621	46.5	320	16	AAK70153
7	1361	39.1	462	22	AAU03646
8	817.5	23.5	727	22	AAU34286
9	817.5	23.5	727	22	AAU37184
10	640	18.4	805	21	AAV75603
11	638	18.3	805	21	AAV75602

12	636	18.3	731	19	AAW44849	S. pneumoniae peni
13	635.5	18.2	805	21	AAV75601	Neisseria gonorrhoe
14	625	17.9	660	18	AAW55545	H. pylori ORF 06ep
15	625	17.9	660	18	AAW55576	H. pylori ORF 06ep
16	608.5	17.5	596	18	AAW89836	Protein encoded by
17	586	16.8	812	22	AAW78604	Lawsonia intracell
18	567	16.3	828	17	AAW04357	E. coli penicillin
19	555	15.9	850	17	AAW86955	E. coli PBP 1A tra
20	552	15.8	774	22	AAU36453	Pseudomonas aerugi
21	538.5	15.5	536	20	AAW89885	Antigen 2 from cin
22	529.5	15.2	821	21	AAV81757	Streptococcus pneu
23	526	15.1	823	17	AAW04358	E. coli penicillin
24	526	15.1	844	22	AAU34446	E. coli cellular p
25	525.5	15.1	781	22	AAU36684	Hemophilus influe
26	519.5	14.9	846	22	AAU38311	Salmonella typhi c
27	514	14.8	844	17	AAW86952	E. coli PBP 1B tra
28	512	14.7	844	17	AAW86953	E. coli PBP 1B tra
29	508	14.6	853	22	AAU55400	Propionibacterium
30	489	14.0	836	17	AAW86954	E. coli PBP 1B tra
31	457	13.1	447	18	AAW20510	H. pylori inner me
32	457	13.1	799	22	AAU49462	Propionibacterium
33	450.5	12.9	770	22	AAW98353	Escherichia coli p
34	441.5	12.7	795	22	AAW90056	C. glutamicum prote
35	437.5	12.6	678	22	AAW81107	Mycobacterium tube
36	422	12.1	720	22	AAW93052	C. glutamicum prote
37	419	12.0	810	22	AAW81232	Mycobacterium tube
38	377	10.8	344	18	AAW55362	H. pylori ORF 14gp
39	377	10.8	344	18	AAW20917	H. pylori cell env
40	370.5	10.6	532	17	AAW86957	E. coli penicillin
41	370.5	10.6	553	17	AAW86956	E. coli penicillin
42	341.5	9.8	231	20	AAW90230	R. eutropha Mgt pa
43	288.5	8.3	269	19	AAW61395	Staphylococcus aur
44	287	8.2	338	22	ABG24982	Novel human diagno
45	280.5	8.1	269	22	AAW81893	S. epidermidis ope

## ALIGNMENTS

RESULT 1  
AAW55063  
ID AAW5063 standard; Protein; 666 AA.  
XX  
XX AAW5063;  
XX  
XX 02-OCT-1998 (first entry)  
XX  
XX Streptococcus pneumoniae sp001 protein.  
DE Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;  
KW Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;  
KW detection; pneumonia; otitis media; meningitis.  
XX  
XX Streptococcus pneumoniae.  
OS  
XX  
XX W09818930-A2.  
PN  
XX 07-MAY-1998.  
PD  
XX 30-OCT-1997; 97WO-US19422.  
XX  
XX 31-OCT-1996; 96US-0029960.  
PR  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Chai GH, Hromockyj A, Johnson LS, Kunsch CA;  
PI  
XX WPI: 1998-272224/24.  
DR  
XX N-PSDB; AAV27323.  
XX  
XX Nucleic acid encoding antigenic peptide(s) from Streptococcus  
PT pneumoniae - or their epitope-containing fragments, useful in  
PR protective or therapeutic vaccines, and for diagnosis

PS Claim 11; Page 48; 118pp; English.

XX The present sequence represents a protein from *Streptococcus pneumoniae*.  
 CC The nucleic acid sequence encoding the *Streptococcus pneumoniae* protein  
 CC can be useful in vaccines for inducing protective antibodies against  
 CC *Streptococcus pneumoniae*, for treatment or prevention of infection e.g.  
 CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid  
 CC are used to detect *Streptococcus pneumoniae* infection (by usual hybridisation or  
 CC amplification methods), also for isolating *Streptococcus pneumoniae* or their  
 CC allelic variants. The protein can be used similarly to detect specific  
 CC antibodies in standard immunoassays, especially for diagnosing or  
 CC monitoring infections. Antibodies which bind the protein are used to  
 CC detect corresponding antigens, to purify the protein and for passive  
 CC immunisation (optionally coupled to a toxin). Vaccines are administered,  
 CC e.g. by injection, orally or through the skin, typically at 0.01-1000  
 CC (especially 10-300)  $\mu$ g/ml per dose.

XX Sequence 666 AA;

Query Match 100.0%; Score 3484; DB 19; Length 666;

Best Local Similarity 100.0%; Pred. No. 1,1e-234; Mismatches 0; Gaps 0;

Matches 666; Conservative 0; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRVNAQANDIPDYLKAIYSIEDHREFFHGRGIDTIRILGAPLRN 60

DB 1 KIYDNKNQIADLGSERRVNAQANDIPDYLKAIYSIEDHREFFHGRGIDTIRILGAPLRN 60

QY 61 LQSNLSGSGSTLTQOLIKLTFSTSTSDQTSRKQAEWALIQLEOKATKQEIILTYINK 120

DB 61 LQSNLSGSGSTLTQOLIKLTFSTSTSDQTSRKQAEWALIQLEOKATKQEIILTYINK 120

QY 121 VYMSNGNYGMQTAONYGKDLNNLSLPOLALAGMPAPNOYDPYSHPEAAODRRNLVL 180

DB 121 VYMSNGNYGMQTAONYGKDLNNLSLPOLALAGMPAPNOYDPYSHPEAAODRRNLVL 180

QY 121 VYMSNGNYGMQTAONYGKDLNNLSLPOLALAGMPAPNOYDPYSHPEAAODRRNLVL 180

DB 121 VYMSNGNYGMQTAONYGKDLNNLSLPOLALAGMPAPNOYDPYSHPEAAODRRNLVL 180

QY 181 SEMKNQYISAQYERKAVNTPITDGLQSLKASANYPAYMDNLYKEVINQVEETGYNLT 240

DB 181 SEMKNQYISAQYERKAVNTPITDGLQSLKASANYPAYMDNLYKEVINQVEETGYNLT 240

QY 241 TGMQVYTVNDDAOKHMDIYNTDEYVAYPPDELOVASTYDVSGKVIQAOLGAAHGSN 300

DB 241 TGMQVYTVNDDAOKHMDIYNTDEYVAYPPDELOVASTYDVSGKVIQAOLGAAHGSN 300

QY 301 VSFQINQAVENRPMGSTMKPTTDYAPALEYGVDSSTATIVHDEPYNPGTWPVYNNDR 360

DB 301 VSFQINQAVENRPMGSTMKPTTDYAPALEYGVDSSTATIVHDEPYNPGTWPVYNNDR 360

QY 361 GYFGNITLQVALQOQRNYPAVETLNKVGILNRAKTFPLNGIDIDPSIHSNMAISSNTTESD 420

DB 361 GYFGNITLQVALQOQRNYPAVETLNKVGILNRAKTFPLNGIDIDPSIHSNMAISSNTTESD 420

QY 421 KKYGASSEKMAAAYAAPANGSTYRKPMITIKHVVSDDGSEKESNNGTRAMKETTAYMMTD 480

DB 421 KKYGASSEKMAAAYAAPANGSTYRKPMITIKHVVSDDGSEKESNNGTRAMKETTAYMMTD 480

QY 481 MKKTLVITGTRNMYLAWLPOAGKTGSNTYDEETENHITKSQFAPELFGAYRKRTSM 540

DB 481 MKKTLVITGTRNMYLAWLPOAGKTGSNTYDEETENHITKSQFAPELFGAYRKRTSM 540

QY 541 AAVTGYSNRLPLVGNGLTVAAYKVRSMKTYLSEGSNPEDNINIPGLRNGEFYKGNAR 600

DB 541 AAVTGYSNRLPLVGNGLTVAAYKVRSMKTYLSEGSNPEDNINIPGLRNGEFYKGNAR 600

QY 601 STWNSPAPQPPSTESSSSSDSSTTPSTNNSTTNNPNNNTQOQNTTPDOQON 660

DB 601 STWNSPAPQPPSTESSSSSDSSTTPSTNNSTTNNPNNNTQOQNTTPDOQON 660

QY 661 POPAOP 666

DB 661 POPAOP 666

RESULT 2

ID AAM04359 standard; Protein: 682 AA.

AC AAM04359;

DT 03-DEC-1996 (first entry)

DE S. pneumoniae penicillin binding protein 1A soluble variant.

KW Penicillin binding protein; PBP 1A; bifunctional protein;

KW transglycosylase; transpeptidase; identification; assay; inhibitor;

KW antibiotic resistant; bacteria; soluble variant; protein structure;

KW X-ray crystallography; determination.

OS Streptococcus pneumoniae.

PN GB2290792-A.

PD 10-JAN-1996.

PE 29-JUN-1995; 95GB-0013306.

PR 24-NOV-1994; 94SE-0004072.

PR 01-JUL-1994; 94IN-0000580.

PA (ASTR ) ASTRA AB.

PI Balganesch TS, Town CM;

DR WP1; 1996-042232/05.

DR N-PSDB; AAT08027.

PS Sol. derivs. of bifunctional penicillin binding protein (BPP) -

PT opt. jack transglycosylase activity, useful to identify and assay

PT for antibodies or cpds. which bind BBPs

PS Claim 2; Pages 65-68; 108pp; English.

XX The present sequence is a soluble variant of the S. pneumoniae

CC penicillin binding protein (PBP) 1A, which lacks the 1st 38 amino

CC acid residues of the wild type protein. Wild type PBP is a

CC bifunctional protein, which binds the cell membrane when expressed

CC in a bacterial cell, having transglycosylase and transpeptidase

CC activities. The variant protein (NCIMB 40665) in conjunction with

CC a labelled anti-bifunctional PBP monoclonal antibody, can be used

CC to identify and assay for cpds. which bind bifunctional PBP. Such

CC cpds., as inhibitors of bifunctional PBP have a potential use in

CC therapeutic cpds. which inhibit the growth of antibiotic resistant

CC bacteria. The soluble variant may also be used in X-ray

CC crystallography.

XX Sequence 682 AA;

Query Match 99.5%; Score 3466; DB 17; Length 682;

Best Local Similarity 99.4%; Pred. No. 2e-233; Mismatches 1; Indels 0; Gaps 0;

Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRVNAQANDIPDYLKAIYSIEDHREFFHGRGIDTIRILGAPLRN 60

DB 1 KIYDNKNQIADLGSERRVNAQANDIPDYLKAIYSIEDHREFFHGRGIDTIRILGAPLRN 60

QY 61 LQSNLSGSGSTLTQOLIKLTFSTSTSDQTSRKQAEWALIQLEOKATKQEIILTYINK 120

DB 61 LQSNLSGSGSTLTQOLIKLTFSTSTSDQTSRKQAEWALIQLEOKATKQEIILTYINK 120

QY 121 VYMSNGNYGMQTAONYGKDLNNLSLPOLALAGMPAPNOYDPYSHPEAAODRRNLVL 180

DB 121 VYMSNGNYGMQTAONYGKDLNNLSLPOLALAGMPAPNOYDPYSHPEAAODRRNLVL 180

QY 181 SEMKNQYISAQYERKAVNTPITDGLQSLKASANYPAYMDNLYKEVINQVEETGYNLT 240

DB 181 SEMKNQYISAQYERKAVNTPITDGLQSLKASANYPAYMDNLYKEVINQVEETGYNLT 240

Db	197	semkrqgylaseqjekavntprltdqglksaansyppaymdyilkevlnqyeeegynllt	258
Qy	241	TGMDVYTVNDQDEAOKHLMDIYNFDEYVAYPPDDELQVASTTVDVNSGKVIAQLGARHGSN	3000
Db	257	tgmdivytnvdgeaqghlwdiynltdeyaypddelqvasltivdangkvilqarhgsn	3168
Qy	301	VSFGTGNQAVETNRBMGSMKRIPTDYAPALEGYVDSTATIVHDEPYNYPGTNPVYVWDR	3606
Db	317	vsfglnqavevtnrwgsmkrlptdyapaleyvyvestaltivhdepynpynlncpvywdr	3766
Qy	361	GYFGNITTLQVALQOOSRNPVAVETLNKVGILNKAFTFLNGLGIDYESIHYSNASISNTESD	4200
Db	377	gyfgnltlqyalqgsrtnpavetlnkvglmkactflngldydpysllysnaisntesd	4368
Qy	421	KRYGASSEKMAAAYAAFAFGSGTYRKPMYIHKVPSDSEKEFSNVGTRAKETAYAMTD	4806
Db	437	krygassekmaayaafangstyykpmylhkvvlsdsgekefsnvgtrankettayamtd	4966
Qy	461	MMKTYLVLTGTGRNAYLAMLPOAGCTGTSNTYDEIEIENHIKTSQVAPDELPACTRKYSM	5400
Db	497	mmktylvltgtrnaylamlpqagktgtsnyldeeiienhiktsqvapdelfaqytrkysm	5566
Qy	541	AVMTGYSTRLLPVLGNGTGLTVAAKYVRSMMTYLSEGSNPEDMNIPEGLYRNGEFVKNGAR	6000
Db	557	avmtgystrllrplvngtlvaakvyrsmtlylsegsnpednlpvglrnygefvkngr	6166
Qy	601	STWNSPAPQDPSTRESSSSSDSSSTSOSSSTTPSTNNTTTNPNNNTQGSNTTPOQONQ	6606
Db	617	stwsnpapqpststresssssdssstsossttpstnntttnpnnntqgsnttpdqngn	6766
Qy	661	PQPAQP 666	
Db	677	pqpapq 682	
RESULT 3			
AAU37830			
ID	AAU37830 standard; Protein: 719 AA.		
XX	AAU37830;		
XX	14-FEB-2002 (first entry)		
XX	Streptococcus pneumoniae cellular proliferation protein #259.		
XX	Antisense: prokaryotic cellular proliferation protein;		
KW	antibiotic; antibacterial; drug design.		
XX	Streptococcus pneumoniae.		
OS	MO200170955-A2.		
XX	27-SEP-2001.		
XX	21-MAR-2001; 2001MO-USO9180.		
XX	21-MAR-2000; 2000US-191078P.		
XX	23-MAY-2000; 2000US-206848P.		
XX	26-MAY-2000; 2000US-207727P.		
XX	23-OCT-2000; 2000US-242578P.		
XX	27-NOV-2000; 2000US-253625P.		
XX	22-DEC-2000; 2000US-257931P.		
XX	16-FEB-2001; 2001US-269308P.		
XX	(ELIT-) ELITRA PHARM INC.		
XX	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;		
XX	Yamamoto RT, Xu HH,		
XX	WPI: 2001-611495/70.		
XX	N-PSDB: AAS55689.		
XX	New polynucleotides for the identification and development of		

PT		antibiotics , comprise sequences of antisense nucleic acids -
XX		
PS		
SS		
Example 3:	Seq ID No 13423:	51pp: English.
CC		The invention relates to antisense inhibitors of genes essential to
CC		prokaryotic cellular proliferation, their use in identifying the
CC		genes themselves and the discovery of novel antibiotics, the essential
CC		genes themselves and the encoded proteins. The prokaryotes used are
CC		Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC		pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC		invention is also useful for the identification of potential new targets
CC		for antibiotic development. The antisense nucleic acids can also be used
CC		to identify proteins used in proliferation, to express these proteins,
CC		and to obtain antibodies capable of binding to the expressed proteins.
CC		The proteins can be used to screen compounds in rational drug discovery
CC		programmes. The antisense nucleic acid sequence is also useful to screen
CC		for homologous nucleic acids which are required for cell proliferation in
CC		a wide variety of organisms. The present sequence represents an
CC		essential prokaryotic cellular proliferation protein.
CC	Note:	The sequence data for this patent did not form part
CC		of the printed specification, but was obtained in electronic
CC		format directly from WIPO at
CC		ftp.wipo.int/pub/published_pct_sequences.
CC		
XX		
SQ	Sequence	719 AA:
Query Match	99.5%;	Score 3466; DB 22; Length 719;
Best Local Similarity	99.4%;	Pred. No. 2.1e-233;
Matches 662:	Conservative 3;	Mismatches 1; Indels 0; Gaps 0.
OY	1 KIYDNKNOLDIADLGSERRVNAQANDIPDVLVAIYAISIEDHREDFRHGIDITRILGAFLRN	60
DB	54 kiydnknqniadiigserrrvnaqandipdlvaivsiedhrfdhrgiditrlilgaflrn	113
OY	61 LOSNSLGSGSTLTQQLIKLTTFSTSTSDOTTSRKQAEMLAIQLEKATKOELITYTNK	120
DB	114 lqsnsllggsalsltlqlkltyfststsdqtlsrkaqaewaiaiqleqkatkgellitylnk	173
OY	121 VYMSNGNNGMORAAONNYVKSDLNNLSLPDLALLAGPQAPNPQYDPYSHPEDAODRRNLVL	180
DB	174 vymnsgnygmqrdaaqnyy9kcdlmmnlslpdlallagmpqapngdyshpeaaqdrnlvl	233
OY	181 SEMKNQGYISAOYEKRAVMTPTTDLGLSLKSASNPAYVDNKLKEVINQVEEETGYNLTT	240
DB	234 semknqgyisaeqykavntpltdglslksasnpyaymdnlylkevinqvveetgynlitt	293
OY	241 TGMADVTVNDQEAOKHLMDIVMTDEVAVPDDELQVASTIVDSNGKVIAQLGARHQSN	300
DB	294 tgmadvtvndgeaokhlwldivmtdevayvpddelqvastlvdsngkvlaqgarhqsn	353
OY	301 VSFGINQAVENTRDMGSTMKPTTDVAPALEGVYSTATIYVDEPYNVNGCTMYPRYNMR	360
DB	354 vsfingnavetrndrgstkpltdvapalegyvastatliydepynpgrtnltpvynmr	413
OY	361 GFYGNTITLOALAQOSRRNVAVETLLNKVGLENRAKTFFLNGIGIDYPSTSHSNAISSWTESD	420
DB	414 gyfgnitltlgylaqgstnrnvavelinkvglnraktfllngigidyprshhsnaissnttesd	473
OY	421 KRYGASSEKMAAAYAFAFANGTTYRKPMYTHKVFSDGSEKESPNVSTRAMKETTYVMMDT	480
DB	474 krygassekmaayaafangelytprmylhkvvfsdgsekesfnvgstramketlaymmdl	533
OY	481 MKKYTLVTGTGNANLAWLMPQAKRGKGTSNYTBDELEHNHTTKSQFYAPAPDELAFGTRYKRISM	540
DB	534 mkkylvltgtgnalawlpdqagktgtsnytbdeleenhltkseqfyapadelafgyctrkrysm	593
OY	541 AAWTGYSNRNLPLVYNGSLTVAAKYVRSMWTYLSSESNPPECDNMIPGLVRNGEFVFRKNAR	600
DB	594 awwtgysnrltplvyngsltvaakyyrsmwtlyssgsnpedcnmipeglvrngefvfrknar	653
OY	601 STWNSPAPQOPSTESSSSSDSTSSGSTTPSTNNSTTTNNPNNNTQQOSTTTPDOQN	660

Db 654 swssppgppstessssdsstgsstpstnstinntpnntqgntirpdpqngn 713  
QY 661 POPAQP 666  
|||||  
Db 714 pqpapq 719

RESULT 4  
ID AAY56106 standard; Protein; 420 AA.  
XX  
AC AAY56106;  
XX  
DT 07-FEB-2000 (first entry)  
XX  
DE Streptococcus pneumoniae pbp1A TER isolate a) protein sequence.  
XX  
KM Streptococcus pneumoniae; penicillin binding protein; pbp2b; pbp1A;  
KW transpeptidase encoding region; TER; antibiotic resistance; diagnosis;  
KM detection; identification; pneumococcal meningitis.  
XX  
OS Streptococcus pneumoniae.  
XX  
PN 2A9807024-A.  
XX  
PD 28-APR-1999.  
XX  
PF 05-AUG-1998; 982A-0007024.  
XX  
PR 01-AUG-1997; 972A-0006886.  
XX  
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.  
PA (UYWI-) UNIT WITWATERSRAND.  
PA (MEDI-) MEDICAL RES COUNCIL.  
XX  
PI Klugman KP, Smith AM, Du Plessis M;  
XX  
DR WPI; 1999-601770/51.  
DR N-PSDB; AAZ35939.  
XX  
PT Polymerase chain reaction assays for detecting Streptococcus pneumonia  
PT useful for the diagnosis of pneumococcal meningitis  
XX  
PS Claim 11; Fig 4; 63pp; English.  
XX  
XX A polymerase chain reaction (PCR) assays have been developed for  
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae  
CC using primers based on the penicillin binding protein 2B (pbp2B) gene  
CC and the pbp1A gene. The products and methods can be used for detecting  
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be  
CC used for simultaneously diagnosing pneumococcal meningitis and  
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.  
CC The methods can be used for detecting S. pneumoniae strains resistant  
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.  
CC The assays can be adapted to detect other pathogens causing meningitis.  
CC S. pneumoniae can be used to detect an antibiotic resistant strain of  
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of  
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and  
CC a 224 bp product. The present sequence represents a Streptococcus  
CC pneumoniae pbp1A transpeptidase encoding region (TER) isolate protein  
CC sequence from the present invention.  
XX  
SQ Sequence 420 AA;

Query Match 63.4%; Score 2208; DB 20; Length 420;  
Best Local Similarity 99.0%; P-Id. No. 6,6e-146;  
Matches 416; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 192 EGYKAVNPTITDGLGSLKASNSPAYMDNYLKEVINQVEETGYNLTTGMQVYTWVQ 251  
|||||  
Db 1 eggkavnpitdglgslksasnypaymdnylkevinqveetgynllttgmqvytlwvq 60

QY 252 EAQKHLMDIYNDEYVAYPDELOAVSTIVDVSNKVIQAOLGARHSSNVSPGINQAVET 311  
|||||  
Db 61 eaqkhlwdiyndeyaypddelqvasctivdsngkvlaqlgarhgsnvsfginqavet 120  
QY 312 NRDKGSTMKPTTDYAPALEYGVDSSTATIVHDEPYNPGTNPVYNNDRGVFGITLQYA 371  
|||||  
Db 121 nrwgsstmkpttdyapalegyvestativhdepyngtncpynwdrqyfgnllqya 180  
QY 372 LOOSRNPVAVETLNKVLINRAKTFNLGLIDYPSIHSNMAISSNTPTSDRKYGASSEKMA 431  
|||||  
Db 181 lqgsrnpvavetlnkvlinraktflngldypsihsnaisntptesdkkygassekma 240  
QY 432 AAYAAFGNGSTYKPKMIIHKVVSDEGSEKESVNGTRAMKETTYAMTDMAKTYLTYGTG 491  
|||||  
Db 241 aayaafangstlykpmiikhvvsdgsesefsnvgtramketaynmtdmktvlsygtg 300  
QY 492 RNMYLAMLPGAGKTGTSNYDEETENIKTSOPAPDELFGATYRKRSMAVMVGSRLT 551  
|||||  
Db 301 rmaylamlpgagktgtsnydeeleniktsqvapdeliaqylrkysmavwvgyrnlt 360  
QY 552 PLVGNGLTVAAKYVRSMMYTLSEGSNPEDMNIPEGLYRNGEFVFKNGARSTWSPAPQP 611  
|||||  
Db 361 plvngnltvaakyrrsmntylsegsnpedwnipeglyrngefvfkngarstwsppqp 420

RESULT 5  
ID AAU35135 standard; Protein; 778 AA.  
XX  
AC AAU35135;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Enterococcus faecalis cellular proliferation protein #422.  
XX  
KW Antisense; prokaryotic cellular proliferation protein;  
KW antibiotic; antibacterial; drug design.  
XX  
OS Enterococcus faecalis.  
XX  
PN WO200170955-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 21-MAR-2001; 2001WO-US09180.  
XX  
PR 21-MAR-2000; 2000US-191078P.  
PR 23-MAY-2000; 2000US-206848P.  
PR 26-MAY-2000; 2000US-207727P.  
PR 23-OCT-2000; 2000US-242578P.  
PR 27-NOV-2000; 2000US-253625P.  
PR 22-DEC-2000; 2000US-257931P.  
PR 16-FEB-2001; 2001US-269308P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;  
XX  
DR WPI; 2001-611495/70.  
DR N-PSDB; AAS52994.  
XX  
PT New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids -  
XX  
PS Example 3; Seq ID No 10728; 511pp; English.  
XX  
XX The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the  
CC genes, their use in the discovery of novel antibiotics, the essential  
CC genes themselves and the encoded proteins. The prokaryotes used are  
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pcl\\_sequences](http://wipo.int/pub/published_pcl_sequences).

Sequence 778 AA:

Query Match 50.3%; Score 1751; DB 22; Length 778;  
Best Local Similarity 51.4%; Pred. No. 1.3e-113;  
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

QY 1 KIYDNKNQIADLGSERRVNAQANDIPDYLKAIYSIEDHREDFHRGIDTIRILGAFRLN 60  
DB 81 klydneifeelgsekreliqndvpgllkdaivsvedirfykhlgydprililgsalsn 140  
QY 61 LOSNLSGGSTLTQOLIKITFYSTSDOTISRKQEMALATOLKOKATKOEILYYINK 120  
DB 141 vkngl1gg9gstlqgl1klystfcsdqclrkqaeamavrlereskeelltyynk 200  
QY 121 VYMSNGNYGMOTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSYSHPEAQRRLNVL 180  
DB 201 vymangfygmetaeenyyghkseldlpqtal1agmpaqansydytkpdtckerrdavl 260  
QY 181 SEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLYKEVINQVEETGYNLT 240  
DB 261 ylmynkkskksaeyekakatpdeg1vp1kaadnrvkvndvynkev1nevkaktgknyvt 320  
QY 241 TGMDDVYTNVDAQKHMDIYNTDEVYAVPPDELOVASTIVYVNSGKYIAOLGARHOSN 300  
DB 321 dgl1dyt1olmdmaqkq1ydvnsdqyva1pddkmgvast1dvaa9gyraq1ygrth1pdd 380  
QY 301 VSFQNGAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGTWPYVNMNR 360  
DB 381 vqlgnllevnlgdvgstevkplmlygpatenlny--stgrlmwdkplkypgcltdidfnsd1 439  
QY 361 GFYGNITLYQALQOSRNPVAVETLNVKGLNRAKTFNLGIDIDPSIHYSNATISNTTESD 420  
DB 440 tygvyltmrtaimgarntc1avq1fdevgken1mpf1kgy1d1ykn1leasna1sntsdvd 499  
QY 421 -KKYGAASSEKMAAAYAAAFANGSTYKKPMYIHKVVFSDGSEKFSNVGRAMKETAYMMT 479  
DB 500 gdkygiissik1laaayaafanng1ynkpyvvnkvfndg1svdpgdqkr1am1ds1aymmt 559  
QY 480 DMKQVTLVYTGGRNAYLAMLPOAGKTGTSNTDEIEENIKTSOPVAPDELPAQTRKYS 539  
DB 560 dmlkvlnvggvgfnga1pg1lgaaktgtsenyded1ar1mgt1ek1lapost1vg1t1y1a 619  
QY 540 MAVMTGYENRILPLVGNGLTYAAKVRSMYTYLSEGSNPEMDNIEGLYRNGEEVFKNCA 599  
DB 620 vsvwtgynldnrlpyg1y1asdy1re1msy1lsgvnsddkvdpdsvrvyn1y1v1kda 679  
QY 600 RSTWN-----SPAPQOPSTESSSSSDSSTQSSTTPTSTNNSTTTNNNNTOOSN 651  
DB 680 yevpr1vqv1p1st1ssap1pess1t1veas1ss1t1keas1ss1ss1s1eap1ss1t1eqpass 739  
QY 652 TTPDOONPOPAP 666  
DB 740 ssaegpatseqp 754

RESULT 6  
AAR70153  
ID AAR70153 standard; Protein: 320 AA.

XX AAR70153;  
AC 14-FEB-1996 (first entry)  
DT Streptococcus pneumoniae strain SPR042 Exp2.  
DE  
XX Exp2; export protein; pbpla; plpa; exp1; exp3; pad1;  
KW virulence determinant; permease like protein;  
KW penicillin binding protein 1A; pyruvate oxidase; regulatory element;  
KW acellular vaccine; antibody.  
OS Streptococcus pneumoniae.  
XX  
XX W09506732-A2.  
XX 09-MAR-1995.  
XX  
XX 01-SEP-1994; 94WO-US09942.  
XX  
XX 01-SEP-1993; 93US-0116541.  
XX 18-MAY-1994; 94US-0245511.  
XX (UYRQ ) UNIV ROCKEFELLER.  
XX  
XX Measure HR, Pearce BJ, Tuomanen E;  
XX WPI; 1995-115448/15.  
XX N-PSDB; AA083241.  
XX  
XX Novel gene fragments encoding specific bacterial exported proteins  
XX - specifically of S. pneumoniae, useful as vaccines  
XX  
XX Claim 35; Page 88-9; 168pp; English.  
XX  
XX This sequence represents exp2. The DNA encoding this sequence is  
XX identical to that for ponA which encodes penicillin-binding protein 1A  
XX (pbpla). This sequence is involved in adhesion of bacteria to target  
XX cells. This sequence is an exported protein of S.pneumoniae. Export  
XX proteins are the proteins in pathogenic bacteria that are virulence  
XX determinants. Other export proteins include pbpa (see AAR70152), exp1,  
XX exp3, and pad1 (encoded by the sequence shown in AA083259). This  
XX sequence can be inserted into an expression vector (preferably a  
XX bacterial expression vector) to provide for high levels of expression of  
XX the protein. The protein can then be used in the production of an  
XX acellular vaccine. These vaccines are used to provide protection from  
XX Gram positive bacterial infection. Antibodies against export proteins  
XX can be used for diagnosis of infection and in passive immune therapy.  
SQ Sequence 320 AA:

Query Match 46.5%; Score 1621; DB 16; Length 320;  
Best Local Similarity 99.7%; Pred. No. 4.3e-105;  
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRVNAQANDIPDYLKAIYSIEDHREDFHRGIDTIRILGAFRLN 60  
DB 5 klydneifeelgsekreliqndvpgllkdaivsvedirfykhlgydprililgsalsn 124  
QY 61 LOSNLSGGSTLTQOLIKITFYSTSDOTISRKQEMALATOLKOKATKOEILYYINK 120  
DB 65 lqsn1lqgssaltqgl1klystfcsdqclrkqaeala1q1ekakqel1tyynk 180  
QY 121 VYMSNGNYGMOTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSYSHPEAQRRLNVL 180  
DB 125 vymangnygm1taeqnyy1gk1dln1slp1q1all1agmpaqansydytkpdtckerrdavl 184  
QY 181 SEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLYKEVINQVEETGYNLT 240  
DB 185 semkngy1saeyekavnlpt1tdg1gs1ksasnpaymdn1y1kev1nqveet1gyn1lt 244  
QY 241 TGMDDVYTNVDAQKHMDIYNTDEVYAVPPDELOVASTIVYVNSGKYIAOLGARHOSN 300

```

Db      245  tgmdivtvdqeaqkhlwdlyntdeyaypddelqvastivdsngkviaglgarhgsn 304
QY      301  VSEGINQAVETNRDNG 316
Db      305  vsfignqaveuicrnwrg 320

RESULT  7
AAU03646
ID      AAU03646 standard; Protein: 462 AA.
XX
AC      AAU03646;
XX
DT      12-SEP-2001 (first entry)
XX
DE      Group B Streptococcus antigenic protein, ID-122.
XX
KW      Group B Streptococcus; encapsulated bacterium; therapeutic; sepsis;
KM      meningitis; neonate; antigenic; vaccine; infection; genital tract;
KM      capsid polysaccharide vaccination.
XX
OS      Streptococcus agalactiae.
XX
PN      WO200132882-A2.
XX
PD      10-MAY-2001.
XX
PF      07-SEP-2000; 2000WO-GH03437.
XX
PR      07-SEP-1999; 99GB-0021125.
XX
PA      (MICR-) MICROBIAL TECHNIQS LTD.
XX
PI      Le Page RMF, Wells JM, Hanniffy SB;
XX
DR      WPI; 2001-316444/33.
XX
DR      N-PSDB; AAS07063.
XX
PT      New polypeptides derived from Streptococcus agalactiae are useful to
PT      provide detection of, and vaccination against, Group B Streptococcus
PT      infections, particularly to prevent infection in neonates -
XX
XX
XX      Claim 1; Fig 1; 178pp; English.
XX
AAU03601-AAU03722 represent Group B Streptococcus (Streptococcus
CC      agalactiae) amino acid sequences of the invention. S. agalactiae is an
CC      encapsulated bacterium which is a major pathogen of humans causing sepsis
CC      and meningitis in neonates as well as adults. The S. agalactiae antigenic
CC      polypeptides are used to vaccinate against Group B Streptococcus
CC      infections, particularly to prevent infection in new born children
CC      arising from the maternal genital tract. An immunogenic composition is
CC      useful in the preparation of a medicament for the treatment of
CC      prophylaxis of Group B Streptococcus infection. The invention does not
CC      have the disadvantages of varied response rate associated with prior art
CC      capsid polysaccharide vaccination against Group B Streptococcus.
XX
XX
SQ      Sequence 462 AA;

Query Match      39.1%; Score 1361; DB 22; Length 462;
Best Local Similarity 62.0%; Pred. NO. 1e-86;
Matches 251; Conservative 73; Mismatches 81; Indels 0; Gaps 0;
QY      2  IYDNKNLIDIGSERVNNQANDIPDVKATISIEDHREFDHRGIDTIRILGAFILRL 61
Db      58  vvggnmliadlqsekresvdsdipnlynatlsiedkxfikrnydyflligaawnhl 117
QY      62  QSNLSIGSTLTQOLIKITFYFSTSDQTSIRKAQAMLAIOLEPOKAKOEIFLYYINKV 121
Db      118  vsnrtqggsrltdqqlklklyfstnksqqltkrksqewlalgmerkytkeellfylnkv 177
QY      122  YMSGNGYGMQTAQNYGYGKDLNLISLPQALLACGMPQAPNOYDEYSHPEAAQDRNLVLS 181

```

```

Db      178  ymgngnygmrttaksyfgkdlkelstlaqllaglpqptqdydknpeseqtrrnevlq 237
QY      182  EMKNQGYISAEQYKAVNPTITDGLQSLKSASNPAVMDNLTKEVINGVEETGYNLTT 241
Db      238  qmygdknlskkeydqavaltvdglkelkqskypkymdnylkqvisvkkqtkgdlfta 297
QY      242  GMDVYTNVNDQEAQKHLMDIYNTDEVAVYPDDELQVASTIVVNSGKYLAQLGARHSSNV 301
Db      298  glkvytnintdaqkqldyindsdyiaypmeqlastimdatnrgkylagqgrhqneml 357
QY      302  SFGINQAVETNRDNGSTMKPITTDYAPALEYGVYDSTATYHDEPYNPGTNPYVNMNRG 361
Db      358  sftngsvltdrdgwtmkpisyayapaldsgvynstgslndsvyypgtatqldwdrg 417
QY      362  YFGNITLOYLALQOSRNPVAVETLNKVGINRAKTFELNGCIDPSI 406
Db      418  ymgwsmqtaiqgsrnpvavraleaagldcaakfleklyglypem 462

RESULT  8
AAU34286
ID      AAU34286 standard; Protein: 727 AA.
XX
AC      AAU34286;
XX
DT      14-FEB-2002 (first entry)
XX
DE      Staphylococcus aureus cellular proliferation protein #562.
XX
KW      Antisense; prokaryotic cellular proliferation protein;
KM      antibiotic; antibacterial; drug design.
XX
OS      Staphylococcus aureus.
XX
PN      WO200170955-A2.
XX
PD      27-SEP-2001.
XX
PF      21-MAR-2001; 2001WO-US09180.
XX
PR      21-MAR-2000; 2000US-191078P.
XX
PR      23-MAY-2000; 2000US-206848P.
XX
PR      26-MAY-2000; 2000US-207727P.
XX
PR      23-OCT-2000; 2000US-242578P.
XX
PR      27-NOV-2000; 2000US-253625P.
XX
PR      22-DEC-2000; 2000US-257931P.
XX
PR      16-FEB-2001; 2001US-269308P.
XX
PA      (ELIT-) ELITRA PHARM INC.
XX
PI      Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI      Yamamoto RT, Xu HH;
XX
DR      WPI; 2001-611495/70.
XX
DR      N-PSDB; AAS52145.
XX
PT      New polynucleotides for the identification and development of
PT      antibiotics, comprise sequences of antisense nucleic acids -
XX
XX
XX      Example 3; Seq ID No 5782; 511pp; English.
XX
PS      The invention relates to antisense inhibitors of genes essential to
PS      prokaryotic cellular proliferation, their use in identifying the
CC      genes, their use in the discovery of novel antibiotics, the essential
CC      genes themselves and the encoded proteins. The prokaryotes used are
CC      Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC      pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC      invention is also useful for the identification of potential new targets
CC      for antibiotic development. The antisense nucleic acids can also be used
CC      to identify proteins used in proliferation, to express these proteins,
CC      and to obtain antibodies capable of binding to the expressed proteins.
CC      The proteins can be used to screen compounds in rational drug discovery

```



**THIS PAGE BLANK (USPTO)**



CLONE: SPRU42  
US-08-600-993A-24

Query Match 11.7%; Score 406; DB 2; Length 77;  
Best Local Similarity 100.0%; Pred. No. 6.2e-24;  
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 240 TTGMVYTNVQEOAKHMDIYNTEDEYVAYPPDELOVASTIYDVSNGKVIQOLGARHOS 299  
Db 1 TTGMVYTNVQEOAKHMDIYNTEDEYVAYPPDELOVASTIYDVSNGKVIQOLGARHOS 60

Qy 300 NVSEGINQAVETNRDMG 316  
Db 61 NVSEGINQAVETNRDMG 77

RESULT 14  
US-08-481-435-12  
Sequence 12, Application US/08481435  
Patent No. 6027906

GENERAL INFORMATION:

APPLICANT: Balganes, Tanjore S

TITLE OF INVENTION: No. 6027906el Polypeptides

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESS: White & Case

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/481,435

FILING DATE: 10-JUL-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: IN 580/MAS/94

FILING DATE: 01-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: SE 9404072-2

FILING DATE: 24-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Sterner, Richard J.

REGISTRATION NUMBER: 35,372

REFERENCE/DOCKET NUMBER: 1103326-151

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 819-8783

TELEFAX: (212) 354-8113

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 532 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Escherichia coli

IMMEDIATE SOURCE:

CLONE: PARC 0593 truncated soluble PBP 1B

US-08-481-435-12

Query Match 10.6%; Score 370.5; DB 3; Length 532;  
Best Local Similarity 30.7%; Pred. No. 6.2e-20;  
Matches 115; Conservative 68; Mismatches 138; Indels 53; Gaps 12;

Qy 9 LIADGSSRRVNAQANDIPTDLVKAIVSIEDHREFDHKGIDTIRILGAFRLNLOS-NSLQ 67  
Db 183 MISSPNCGRFLFVPRSGFPDLVDLTATEDRHFYEHGDISLYSIGRAVLNFTAGRTVO 242

Qy 68 GGSITLQOLIKLTYSTSTSDQITSRKAQEAFLAQLQOKATKQIILYINKYV---- 123  
Db 243 GASTLTQOLVKNLFLS---SERSVWRKANEAVMALIMARYSKDIILLYNVEVYLGOSG 299

Qy 124 SNGVGMCTAQNQYKQDLNLSLPOLLALAGMPQAPNOYDPYSHPEAAOORNLVSEM 183  
Db 300 DNEIRGFPLASLYTGRVYEELSLDQALVGMVKGASITVNPWRPKLALERNLVLL 359

Qy 184 KNQYISAEOYERKAVNTPITDLOSLKSA-SNYPAYMDNLYKEVI---NOVEEFTGYNL 238  
Db 360 QQQQIILDELDYDMLSARPL--GVQPRGVITSPQAPFMQLVNGELQAKGLDKYKDLGVKI 417

Qy 239 LTTGMVYTNVQEOA-----QKHMDIYNTEDEYVAYPPDELOVASTIYVS 284  
Db 418 FTT---FDVVAQDAEAKAAGEGIPALKKORLSD-----LETAIVVYDRF 459

Qy 285 NGKVIQOLGARHOSNVSEF-GINQAVETNRDMGSTMKPITDYAPALEYGVYDSTATVHD 343  
Db 460 SGVRAAWG----GSEPPAGINRAMQARRSIGSLAKPAT-YLRALSQPKIYRLNTWIAD 514

Qy 344 EPYNYPGTNTPYN 357  
Db 515 APIALROPNGQVWS 528

RESULT 15  
US-08-481-435-11  
Sequence 11, Application US/08481435  
Patent No. 6027906

GENERAL INFORMATION:

APPLICANT: Balganes, Tanjore S

TITLE OF INVENTION: No. 6027906el Polypeptides

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESS: White & Case

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/481,435

FILING DATE: 10-JUL-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: IN 580/MAS/94

FILING DATE: 01-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: SE 9404072-2

FILING DATE: 24-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Sterner, Richard J.

REGISTRATION NUMBER: 35,372

REFERENCE/DOCKET NUMBER: 1103326-151

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 819-8783

TELEFAX: (212) 354-8113

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 553 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

```

Db 529 PIALROPNGVSPQNDRRYSESGRMLVDALTRSMNPTVLGMLGLPAVETWIKL 588
QY 400 GIDYPSIHNSNAI---SSNTESDCKKYGASSEKMAAAYAFANGCTYYKMYIHKVYFSD 456
Db 589 GVPKDLHPVPAULGALNLTPIE-----VAAQFOTIAGGNAPLSALRSVIAED 639
QY 457 GS--EREFSNVGIRAKETAYAMMTDMKVTLTGTGR--NAVIALPQAGKTGTSNYTD 512
Db 640 GKLYGSFQ-ARAVPAQAAYLTLMTMOYVGRGTCRGAGKAYPNLHLAKGTGTN--- 695
QY 513 EELENIHKISQVAPDELPAGYTRKYSMAVWICYSNRLPLVGNGLTVAAKYRSMYTL 572
Db 696 ---NNV-----DTWFAIGDSTVTTITWGRDNN-OPTKLYGASGAMSIYQ---RYL 739
QY 573 SEGSDNEDNII--PE-----GLYRNGEFPVKNGAR---STWNSPAPQOPPTSESSSSSDS 623
Db 740 A-NQPTPLVLPEDIDADMVDYDGNFVCGSGMRLLPVWTS-----DPQSLCQGS 789
QY 624 STSQSSSTPSTNNTTTPNNTTQOQNTTPOQONPOPAQ 665
Db 790 EMQ-----QPSGNPFOSSQPOQPOQ-QPAQ 816

RESULT 12
US-08-245-511-24
; Sequence 24, Application US/08245511
; Patent No. 5928900
; GENERAL INFORMATION:
; APPLICANT: Measure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6

```

```

; IMMEDIATE SOURCE:
; CLONE: SP042
; US-08-245-511-24

Query Match 11.7%; Score 406; DB 2; Length 77;
Best local Similarity 100.0%; Pred. No. 6.2e-24;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 TTGMDYTVNDQAKHMDIYTFDEVAVPDELOVASTIVYVSGKYAQLGARHSS 299
Db 1 TTGMDYTVNDQAKHMDIYNTDEVAVAPDDELQVASTIVYVSGKYAQLGARHSS 60
QY 300 NVSEGINQAVETNRDNG 316
Db 61 NVSEGINQAVETNRDNG 77

RESULT 13
US-08-600-993A-24
; Sequence 24, Application US/08600993A
; Patent No. 5961229
; GENERAL INFORMATION:
; APPLICANT: Measure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6
; IMMEDIATE SOURCE:

```

```

?      TYPE: amino acid
?      STRANDEDNESS:
?      TOPOLOGY: linear
?      MOLECULE TYPE: peptide
?      ORIGINAL SOURCE:
?      ORGANISM: Escherichia coli
?      IMMEDIATE SOURCE:
?      CLONE: pARC0468 PBP 1B QOQL
US-08-481-435-8

```

Query Match	14.7%	Score 512;	DB 3;	Length 844;
Best Local Similarity	28.6%	Pred. No. 2.5e-30;		
Matches 201;	Conservative 99;	Mismatches 275;	Indels 128;	Gaps 30;

QY	9	LIADLGSRRVNAOANIPDTLVVAIVSIEDHREDFHRGIDTRIGLAFGRNLOS--NSLD	67
Db	204	MISSPNNGQRFLVFRSGFPRDLVDTLLATEDRHHYEDHDSLSXISGAVALANTLACTGYQ	263
QY	68	GGSTLTQOOLIKLYFSTSTSDQITSRAQEAMLAIQLEOKATQOEILITYINKVM---123	305
Db	264	GASTLTLLVKNLELS---SERSYWRKARNEAYMLIMADARYSDRIILELYMNEYVIGOSG	320
QY	124	SNGVNGQTAQONRYGKDIINLSIPOLALLAGMOPARNOVDYPSHPPAADRRNNIVISEM	183
Db	321	DNEIRGFPLASLYFFGRVEELSLDOOQLVGVKAGASTYPMNPKLALERRNLVLRLL	360
QY	184	KNQGYISAQYKAVNTPITDGLQSLKSA--SNPAYVDNLYKEVI-----NQVEETGYNL	238
Db	381	QOQOQIIDDELMDLMSARPL--GVQPRGVGISPPQAFMQVLQROELQAKLGDKVADLSGVKI	438
QY	239	LTTGMDVYTNNDQA-----OKHMDIYNTDEYAVPPDELOQVASTIVDS	284
Db	439	FTT---FDSYAQDAEAKAAYEGIPALKKQKSLD-----LETAIYVVDPR	480
QY	285	NGKYIAQLGARHOSNSNF--GINQAVETNRDMGSTMKPIPTYAPALEGYVDSTATIVHD	343
Db	481	SGEVRAMG---GSEQFQAGYNNAQARISIGSLAPAI--YLTALLSQPIITYKLNTMIAD	555
QY	344	EPLYNPGTNPVY---NMDRGY--FGNITLOALQOSRNPAYVETLKNKVLNRAKTFPLNG	398
Db	536	APILRQPNCGVMSPOPNDRRYSESGRMIVYDALTRSMNPVYMLGNALGIPAVETWIK	595
QY	399	LGIDVPSIHYSNAL---SNTTESDKKYVASSEKMAAAYAFANCGYKYKPYIHKVYFS	455
Db	596	LGVRKDDQHLPRPALGLALNTPLE-----VAAQFOTIASGGRAPLSALRSVIAE	646
QY	456	DGS--EKEESNVGTRAMKETTAYMYTDMKTKVLTYGTR--NAYLAWLPOAGKTGTSNYT	511
Db	647	DGKVLVYOSFPQ--AERAVAPQAAYLTLLTMQOVYQGRDOLGAKKYPMLJHLAGTGTTN--	703
QY	512	DEBIEHNHKTSQFAVAPBELFAGYTRKXSMVNGYSNRLLPVLVNGGLTVAAKYRSMY	571
Db	704	-----NNV-----DTWIFAGIDGCTVYTLTWGRDNN--OPTKLYGSGAMSTIYQ--RY	746
QY	572	LSGGSNPEDMNI--PE-----GLYRNGEYFKNGAR--STWNSFPAPQOPSTSTSSSSSD	622
Db	747	LA--NOTPLPLVLVPEDIDAGVDGYDDGJFVCSGGMRILLPYWTS-----DPQSLCQ	796
QY	623	SSTGSSSSTPTSTNNSTTNNNNTQOASNTPPDQOONOPQAPQ	665
Db	797	SEMOO-----QSGNPFDOSSOPQOOPQO	824

```

RESULT 11
US-08-481-435-9
: Sequence 9, Application US/08481435
: Patent No. 6027906
: GENERAL INFORMATION:
: APPLICANT: Balganes, Tanjore S
: APPLICANT: Town, Christine
: TITLE OF INVENTION: No. 6027906el Polypeptides
: NUMBER OF SEQUENCES: 42

```

CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036-2787  
  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,435  
FILING DATE: 10-JUL-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IN 580/MAS/94  
FILING DATE: 01-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9404072-2  
FILING DATE: 24-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sterner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-151  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 836 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli  
IMMEDIATE SOURCE:  
CLONE: PARC0469 PBP 1B del 8  
  
US-08-481-435-9

[illegible]

RESULT 9  
US-08-481-435-7  
Sequence 7, Application US/08481435  
Patent No. 6027906  
GENERAL INFORMATION:  
APPLICANT: Balganes, Tanjore S  
APPLICANT: Town, Christine  
TITLE OF INVENTION: No. 6027906el Polypeptides  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,435  
FILING DATE: 10-JUL-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IN 580/MAS/94  
FILING DATE: 01-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9404072-2  
FILING DATE: 24-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Steiner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-151  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 844 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli  
IMMEDIATE SOURCE:  
CLONE: PARC0438 PBP 1B CQAA  
US-08-481-435-7

Query Match 14.8%; Score 514; DB 3; Length 844;  
Best Local Similarity 28.6%; Pred. No. 1.8e-30;  
Matches 201; Conservative 99; Mismatches 275; Indels 128; Gaps 30;

QY 9 L1ADLGSERRVNAQANDIPTDLVKAIVSTEDHFFPDHGRGIDTIRILGAPLRLNQS-NSLQ 67  
Db 204 MISSPENGEOPLVPRSGFPLDLVTLATEDRHFYEHGDISLXISIRAVLANITAGRTVO 263  
QY 68 GSGTLTQOLIKLTYFSTSDTISRKAQEWLAIOLEOKATKOELITTYINKVYV---- 123  
Db 264 GASTLALVKNFLS-----SERSYWRKANAYVALLIMDARYSKDRILLELMNRYVIGQSS 320  
QY 124 SNGNGMGTAAONYGKDLNNISLPOLALLAGMPQAPNOYDPSHPEAADRNLVLSQM 183  
Db 321 DNEIRGFPLASYFYEGRPVELSLDQALLVGMVKGASITYNPRNPKLALERRNLVLRLL 380  
QY 184 KNGYISAQYEAQVPTPTDGLQSLKSA-SNPATMDNYLKEVT-----NOVEEELGYNL 238  
Db 381 0000IIDOLYDLMLSLARPL--GVQPRGCVISIPQAFMOLVRELQAKLDKDKVSLGVKI 438  
QY 239 LFTGMDEVYTNVQEA-----QKHLMDIYNDEYVAYVDDDELQVASTIVDS 284

Db 439 FTR-----FDSVADDAEKAKEAVEGIPALKKQKRLSD-----LETAIVVDFR 480  
QY 285 NGKVIQALGARHOSNVSF-GIQAVETNNDKSGTAKPTIDYAPALEYGYDSTATIVHD 343  
Db 481 SGYERAMVY-----GSEPOFAGYNRAQARRSIGSLAKPAT-YLTALSQPKIYRLNWIAD 535  
QY 344 EPNYVPGTNTPPVY-----NWDGKY--FGNITLOVALOOSRNPVAVETLNLKVLNRAKTFLLNG 398  
Db 536 APLALQPNQGVWSPQNDDBRYSESGRVMLVDALTSMMVPTVNLGMALGLPAVTEWIK 555  
QY 399 LGIDYPSIHYSNAI-----SSNTTESDKKYGASSEKMAAAYAFANGGTYYKPMYIHKVFS 455  
Db 596 LGVPKQQLHPVPMALGLNLFTPE-----VAQAFQTIASGGRAPLALSRIYAE 646  
QY 456 DGS--EKESNVGTFRAMKETTATMTDMAKTUTLYTGCR--NAYLMLPQAGTGTSNNT 511  
Db 647 DGVLYQSFPO-AERAVPAQAAVLTMTMOQVVGRTGROLGAKYPNLHLAGKRTGTTN-- 703  
QY 512 DELEENHIKTSQFVAPDELPAQYTRKYSMAVMTGYSNRLPLVGNGLTVAAKYRSMTY 571  
Db 704 -----NIV-----DYMFAGIDGSTVITTWVGRDNN-OPTKLYGASGAMSTIYQ---RY 746  
QY 572 LSEGSNPEDMNT--PE-----GLYRGEFVFKNGAR--STWNSPAPQOPPTSESSSSSD 622  
Db 747 LA-NQPTPLNLVPEPIDADMGVDYDGNFVCSGMRILPWTS-----DPQSLCQ 796  
QY 623 SSTSSQSSSTPTSTNNSTTTPNNNTQOQSNTPPQOQNOQPAQ 665  
Db 797 SEMQO-----QPSGNPQDSSQPOQOPOQO-QPAQ 824

RESULT 10  
US-08-481-435-8  
Sequence 8, Application US/08481435  
Patent No. 6027906  
GENERAL INFORMATION:  
APPLICANT: Balganes, Tanjore S  
APPLICANT: Town, Christine  
TITLE OF INVENTION: No. 6027906el Polypeptides  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,435  
FILING DATE: 10-JUL-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IN 580/MAS/94  
FILING DATE: 01-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9404072-2  
FILING DATE: 24-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Steiner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-151  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 844 amino acids

Oy	237	NULLTGMDVYNNVNO	OEAO-----KH-----LMDIYN-----TDEYYA	268
Db	286	RIYTT---ITRKVQAAAO	QAVRNNVLDYDMRHGTRGPANLMMKYGESAMDNKITTDLKA	342
Oy	269	YP-----	-----DDEL	274
Db	343	LPYTGPIIPAAVTSAN	POQATAMLADSGTVALMSMEGRMARPRYRSDTQOGPTPRKRYTDVL	402
Oy	275	Q-----	-----VASTIYDVS--NGXIYAOLGA--RHQSSNVSEGLNOAV	309
Db	403	QTGQGIWVROYGDAMW	LAQYPEVNSALVTSINPQGAVMALVGGDFNQSK-----FNRRAT	457
Oy	310	ETNBDGSTMKPTINDY	PALEIEGYDSTATIHHDEPNYNGTNPVYNNMDRG-----	361
Db	458	QALROVGSNIPEFL-TR	AMNKGKL--TLASMLND-----VPISRMASAGSDMOPK	505
Oy	362	-----YFENITLQYAL	ODOSRRNVPVETLNNKGYLMRAKTFNLGLIDYPSIHSNAISSNT	416
Db	506	NSPQYAGPIRLRGGL	QSKNVVWVRAMRAMGVYDAEYLQRFQFPQNLTVHTESLA---	562
Oy	417	TESDKTKGASS---	EKMAAAVAAAPANGCTYTKPMYIHKVYFSDG-----	457
Db	563	-----LGSASFPTMOY	ARGVAVANANGFLVDPWFIFISKIENDQGGVYFEAKPRVACECD	616
Oy	458	-----SEKEFSNV-----	CTRAMKETTAYM--MTDMKATVLTGTGRNAY-----	495
Db	617	IPVYIGDTQKSNLEN	DVEDVALISEQONVSVPMPOLEQANQALVAKTAQAEAPRVIN	676
Oy	496	---LAMLPOA-----	-----GRTGTSNYDEEIEHNHIKTSQF	524
Db	677	TPLAFLIKSALNTN	IFGEPGQGTGWRAGRDLQDRIDGKTGTN-----	722
Oy	525	VAPDELPGATYRKRS	MAVMYTSYSHRLPIPLYGNGLTVAAKYRRSMYTLSEGSNPEDMNIP	584
Db	722	SSKAWPESGYCPGV	YTSWITGFIQDHRNRL--GHTTASGAIKDQISYEBSGA-----	770
Oy	585	EGLYRNGEFYFKNAR	STWNS-----PAPQOP--STSESSSSSDSTSSQSSSTPST	635
Db	771	-----KSAQPMQWIM	YMAVLEGEVPEOPLTPPGIIVYNIDRSTGO-----	810
Oy	636	NNSTTTPNNNTQO--	SNTTDPQO	657
Db	811	----LANGNSREYEF	TGCTPQO	831
RESULT	8			
US-08-481-435-4				
Sequence 4, Application US/08481435				
Patent NO. 6027906				
GENERAL INFORMATION:				
APPLICANT: Balanesh, Tanjore S				
APPLICANT: Town, Christine				
TITLE OF INVENTION: No. 6027906el Polypeptides				
NUMBER OF SEQUENCES: 42				
CORRESPONDENCE ADDRESS:				
ADDRESSEE: White & Case				
STREET: 1155 Avenue of the Americas				
CITY: New York				
STATE: New York				
COUNTRY: United States				
ZIP: 10036-2787				
COMPUTER READABLE FORM:				
MEDIUM TYPE: Floppy disk				
COMPUTER: IBM PC compatible				
OPERATING SYSTEM: PC-DOS/MS-DOS				
SOFTWARE: PatentIn Release #1.0, Version #1.30				
CURRENT APPLICATION DATA:				
APPLICATION NUMBER: US/08/481,435				
FILING DATE: 10-JUL-1995				
CLASSIFICATION: 435				
PRIOR APPLICATION DATA:				
APPLICATION NUMBER: IN 580/MAS/94				

```

1      FILING DATE: 01-JUL-1994
2      PRIOR APPLICATION DATA:
3      APPLICATION NUMBER: SE 9404072-2
4      FILING DATE: 24-NOV-1994
5      ATTORNEY/AGENT INFORMATION:
6      NAME: Sterner, Richard J.
7      REGISTRATION NUMBER: 35,372
8      REFERENCE/DOCKET NUMBER: 1103326-151
9      TELECOMMUNICATION INFORMATION:
10     TELEPHONE: (212) 819-8783
11     TELEFAX: (212) 354-8113
12     INFORMATION FOR SEO ID NO: 4:
13     SEQUENCE CHARACTERISTICS:
14     LENGTH: 823 amino acids
15     TYPE: amino acid
16     TOPOLOGY: linear
17     MOLECULE TYPE: protein
18     OS-08-461-435-4

```

Query Match	15.1%;	Score 526;	DB 3;	Length 823;
Best Local Similarity	28.9%;	Pred. NO. 2.1e-31;		
Matches 203; Conservative	99;	Mismatches 273;	Indels 128;	Gaps 30

QY	9	LIADLGSRRVAAOANDIPTLVAAIVSIEDHREDFHJRGDITILIAFLFNLOS	67
Db	183	MISSPNEGQRFLVFRSGPDLVLVOTLATBDRNHVEYHNDGISIVSIGRAVLANTLACTGYQ	242
QY	68	GGSTLLQOOLKLYFSTSDQTSIKRAQEAALIOLEOKATKOELLITYINKVM	123
Db	243	GASTLLQOOLVKNELFS	299
QY	124	SNGVYCGMOTAAONVYCGKOLNLSIPOLALLAGMPOARNOVDPYSHPEAODRRLVISEM	183
Db	300	DNEIRGFPLASLYFFGRVEELSLDOALLGVYWKAGSITYPMPNPKLALERRNLVIRLL	355
QY	184	KNQGYISAEQYKAVNTPITDGLQSLKSA	238
Db	360	QOQOOLIDELVDMLSARPL	417
QY	239	LTTGMVDYTNDOEA	284
Db	418	FTT	455
QY	285	NGKYIALQAGARHOSSNVSF	343
Db	460	SGEVRAWVG	514
QY	344	EPPYNPCTNTPVY	398
Db	515	APILARQNGVWSPONDNRYSSESGRMVLVDALTRSNVPTVMLGHALGIPAVETWIK	574
QY	399	LGIDYPSIHYSNAL	455
Db	575	LGVRKDDQLHPAPMLLALNLTPIE	623
QY	456	DGS	511
Db	626	DGKLYLOSFPQ	682
QY	512	DEELIENHIKTSQPAAPDELFAGYTRKYSMAWVGYSNRLPIVNGSLTVAAKYRSMATY	571
Db	683	NNV	725
QY	572	LSGSENPEDMNI	622
Db	726	LA	775
QY	623	SSTSQSSSTPSTNNSTTNNNTQOOSNTTPDQONONOPAQ	665
Db	776	SEMOO	803

TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 828 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-481-435-2

Query Match 16.3%; Score 567; DB 3; Length 828;  
Best Local Similarity 23.4%; Pred. No. 1.7e-34;  
Matches 202; Conservative 129; Mismatches 244; Indels 290; Gaps 30;

QY 1 KIYDNKNQIADLGSERRRANQAANDIPTDLVKAIVSIEDHREFDHRGIDTIRIL-GAFLR 59  
DB 27 QIYSADGELIAOYGERRRIPVTLDOIPEPMKAFIATEDSRFEYEHNGVDPVGIFFRAASVA 86  
QY 60 NIQNSLGGSTLTQOLIKITYFSTSDQTSRKAQEAAMLAIOLEOKATKOEILTYIYN 119  
DB 87 LFSGHSQCASTITQOLANRFLS---PERTLMRKIKEVFLAIRIQLLTKEILELYLN 143  
QY 120 KYVMSNGNYGMOTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPEAAQDRRLV 179  
DB 144 KIYLGFRAYGVGAAGVYFGKTVDOQLTNEMAVIAGLPKAPSTFENFLYSMDRAVARRVV 203  
QY 180 LSEMKNOGIYSAQYERKAVNTPITDGLQSLKSASNPAYMDNLKVINQVEE---ETGY 236  
DB 204 LSRMDEGIYTOOQFDQTEAIANNYHAFELAFSAPIYSEVNRQMYNRKYGSAVEDGY 263  
QY 237 NLTGMDVYTVNDQEAQ-----KH-----LMDIYV-----TDEYVA 268  
DB 264 RIYTT---ITRRVQQAQAQAVNRNNVLDYDMRHCRGPANVLMKVGSAANDNKITTLKA 320  
QY 269 YP-----ITRRVQQAQAQAVNRNNVLDYDMRHCRGPANVLMKVGSAANDNKITTLKA 320  
DB 321 LPTYPGLPAAVTSANPQOATAMLADGSTVALSMESGVRWAPRSDTQOGPTPRKVTDLV 380  
QY 275 Q-----VASTIYDVS--NGKVINQLGA--RHQSSNSFGLNOAV 309  
DB 381 QYGGQIIVWROVDAMMLAOPVENSAIVSINPQNGAVMLVGGFDPNQSK---FNRAI 435  
QY 310 ETNRDQSTMKPITDYPALALEYGVYSTATIVHDEPYNPGTTPYNNDRG----- 361  
DB 436 QALROVGSNIKPF--YTAAMDGL--TLASMLND-----VPSLRDASAGSWMOK 483  
QY 362 ----YFGNITLQYALQGSRNPAVETLTKVGLNRKATFLNGLGIDYPSIHVSNAISNT 416  
DB 484 NSPPQYAGPIRLRQGLGQSKNVYVVRAMRAMGVDAEYLQRFQPAONIVHTESLA--- 540  
QY 417 TESDKKYGAS---EKMAAAYAFANGGTYYKPMYIHKAVFSDG----- 457  
DB 541 ----LGSASFTPMQVANGVAVMANGFLVDPMFISKIENDGGVIFPAKRVACPECD 594  
QY 458 ----SEKEFSNV-----GTRAMKETTAYM--MTDMKTVLYLGGGRNAY----- 495  
DB 595 IYVYIGDQTKSNLVNFGEPGMOGTGWRAGRDQLQRDIGKLTGTN----- 699  
QY 496 --LAWLPOA-----GKTGTSNYTDEIEIENHIKTSOF 524  
DB 655 TPLAFLIKALMTNIFGEPGMOGTGWRAGRDQLQRDIGKLTGTN----- 699  
QY 525 VAPDELFCAGTYKSKAVVTGYSNRILTPVNGGLVYAAVYVSMYTYLSEGSNPEDWNIP 584  
DB 700 SSKDAMFSGYGVGVYVSWIGFDHRRNL--GHTTASGAIDQISGYEGCA----- 748  
QY 585 ECLYRNGEVEFNKAGASTNS-----PAPOP--PTSESSSSSSDSTSOSSTPTST 635  
DB 749 -----KSAQAPANDAYMKAVLEGVPEQPLTPPGIYIVNIDRSTGQ----- 788  
QY 636 NNSITTNPNNTQO---SNTTDDQ 657  
DB 789 ----LANGNSRREYFIETGPTQO 809

## RESULT 7

US-08-481-435-10

Sequence 10, Application US/08481435

Patent No. 6027906

GENERAL INFORMATION:

APPLICANT: Balganes, Tanjore S

APPLICANT: Town, Christine

TITLE OF INVENTION: No. 6027906el Polypeptides

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: White &amp; Case

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/481,435

FILING DATE: 10-JUL-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: IN 580/MAS/94

FILING DATE: 01-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: SE 9404072-2

FILING DATE: 24-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Steiner, Richard J.

REGISTRATION NUMBER: 35,372

REFERENCE/DOCKET NUMBER: 1103326-151

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 819-8783

TELEFAX: (212) 354-8113

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 850 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Escherichia coli

IMMEDIATE SOURCE:

CLONE: PARC0571 BFP 1A Q0AA

US-08-481-435-10

Query Match 15.9%; Score 555; DB 3; Length 850;  
Best Local Similarity 23.1%; Pred. No. 1.4e-33;  
Matches 200; Conservative 129; Mismatches 246; Indels 290; Gaps 30;

QY 1 KIYDNKNQIADLGSERRRANQAANDIPTDLVKAIVSIEDHREFDHRGIDTIRIL-GAFLR 59  
DB 49 QIYSADGELIAOYGERRRIPVTLDOIPEPMKAFIATEDSRFEYEHNGVDPVGIFFRAASVA 108  
QY 60 NIQNSLGGSTLTQOLIKITYFSTSDQTSRKAQEAAMLAIOLEOKATKOEILTYIYN 119  
DB 109 LFSGHSQCASTITQOLANRFLS---PERTLMRKIKEVFLAIRIQLLTKEILELYLN 165  
QY 120 KYVMSNGNYGMOTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPEAAQDRRLV 179  
DB 166 KIYLGFRAYGVGAAGVYFGKTVDOQLTNEMAVIAGLPKAPSTFENFLYSMDRAVARRVV 225  
QY 180 LSEMKNOGIYSAQYERKAVNTPITDGLQSLKSASNPAYMDNLKVINQVEE---ETGY 236  
DB 226 LSRMDEGIYTOOQFDQTEAIANNYHAFELAFSAPIYSEVNRQMYNRKYGSAVEDGY 285



D <b>b</b>	185	SEMNOCIGISAEOXERKANVFPTGLOSLKSASNYPAYMNYLKEVHNOYVEEFGYHLTT	244
O <b>y</b>	241	TGMDVYTVWDD EAO KHLMDIYNTEDEYAAVPDD ELOVASTIVDSNGKVINOLGARHOSSN	300
D <b>b</b>	245	TGMVYTVWDD EAO KHLMDIYNTEDEYAAVPDD ELOVASTIVDSNGKVINOLGARHOSSN	304
O <b>y</b>	301	VSEGINOA VETNRDMG 316	
D <b>b</b>	305	VSEGINOA VETNRDMG 320	

RESULT 5  
US-08-731-

Sequence 2, 5789202  
Patent No. 5789202  
GENERAL INFORMATION:  
APPLICANT: Hoskins, JoAnn  
APPLICANT: Jaskunas, S. Richard  
APPLICANT: Rokey, Pamela K.  
APPLICANT: Zhao, Genshi  
APPLICANT: Kostec, Paul R. Jr.  
TITLE OF INVENTION: penicillin Binding Protein From  
TITLE OF INVENTION: Streptococcus Pneumoniae  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/731,716  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39, 872  
REFERENCE/DOCKET NUMBER: X-10, 887  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 731 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-731-716-2

Query Match	18.3%	Score 636	DB 1	Length 731
Best Local Similarity	27.7%	Pred. No. 8.2e+40		
Matches 192	Conservative 108	Mismatches 252	Indels 142	Gaps 19

```

Db      269  NMVAAGYIDKNOEHEAAVEVDMTSQLDHHOKYEBEKISDYRTPSTFPAVVAEASKNLTDEE- 3227
      | | | | | : : : | | | | | : : : | | | | |
Qy      225  GYNLLTTCGMADVITNDVDEQAOKHMLDIY-NTDEYVAYPDDEL-QVASTIVDVSNGKYTAOL 2922
      : : : : : | : : : | : : : | : : : | : : : | : : : |
Db      328  ---IVNNGYRIYTELDQNYQANMOIYVENMTSLFPRADGTFFAOGSGVALBPKTGCVAGVY 3844
      : : : : : | : : : | : : : | : : : | : : : | : : : |
Qy      293  GARHOSSVNSF-GINQVAEYTRMDGSGTMRPTTDAFPALEKYGVYDSTATIYHDEPY----- 3466
      | | | | | : : : | | | | | : : : | | | | |
Db      385  GOVANDNDKGTGRNENYATQSKRSRSGSTIKPLVYTPFAVEGGMALNKOGLDHTMOYDSYKV 4444
      : : : : : | : : : | : : : | : : : | : : : | : : : |
Qy      347  -NYPGTNT---PYVMNDRCYFGNITLQYALQOOSRNPAYEVELTNKYGLNRAKTFPLNGIGI 4010
      | | | | | : : : | : : : | : : : | : : : | : : : |
Db      445  DNYAGIKTSREVPYKQ-----SLAESLNPAYATVNDLSVDYA----- 4422
      : : : : : | : : : | : : : | : : : | : : : | : : : |
Qy      402  DYPSTHYSNAISSMTTESDKKYGASSER-----MAAAYAFAFANGT 4422
      : : : : : | : : : | : : : | : : : | : : : | : : : |
Db      483  -----PEAGEKFGELMKRKYDRLVGLGVALGSGVETNPLOMAQAYAAFAANEGL 5277
      : : : : : | : : : | : : : | : : : | : : : | : : : |
Qy      443  YKKPMYITHKYVFSQGSF-KEPSNNGTSAAMEKTTAYMMTDMMKITYLYTGTRNAYLAMPQ 5010
      : : : : : | : : : | : : : | : : : | : : : | : : : |
Db      528  MPEAHFISIRINAGGOYIASHKNSQOKRYADIKSVADKMTSMNLGTFTGTGSISSPAAUYM 5877
      : : : : : | : : : | : : : | : : : | : : : | : : : |
Qy      502  AGKGTGTSYNTDEELENHKTQSOPAPDELPAGYTRYKYSMAWNTGY-----SNRLTPLYVGN 5566
      | | | | | : : : | | | | | : : : | | | | |
Db      588  AGKGTGTT---EAYFNEPYTS-----DQWYIGYTPDVYISHMLGFTPTDENHLYAGTSN 6388
      : : : : : | : : : | : : : | : : : | : : : | : : : |
Qy      557  GLTVAAKYVRSMMTYLSEGSNPEEDMNTPEGLYRNGEEVFKNGARSTWNSPAPQOPSTES 6166
      | | | | | : : : | | | | | : : : | | | | |
Db      639  G---AAHFERNIAMTIL-----PYTPG 6577
      : : : : : | : : : | : : : | : : : | : : : | : : : |
Qy      617  SSSSDSTSGSSSTSTBTSTNNSTTTNNNNNTQOS 650
      : : : : : | : : : | : : : | : : : | : : : | : : : |
Db      658  STFTVENAYKONGIAPATWKROYGTNNNSQTDN 651
      : : : : : | : : : | : : : | : : : | : : : | : : : |

```

RESULT 6  
US-08-481-435-2  
Sequence 2, Application US/08481435  
Patent No. 6027906  
GENERAL INFORMATION:  
APPLICANT: Balanesh, Tanjore S  
APPLICANT: Town, Christine  
TITLE OF INVENTION: No. 6027906el Polypeptides  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,435  
FILING DATE: 10-JUL-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IN 580/MNS/94  
FILING DATE: 01-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9404072-2  
FILING DATE: 24-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Stener, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-151  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783

```
RESULT 3
US-08-245-511-4
; Sequence 4, Application US/08245511
; Patent No. 5928900
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 343-1684
; TELEFAX: 201 487-5800
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-245-511-4

Query Match 46.5%; Score 1621; DB 2; Length 320;
Best Local Similarity 99.7%; Pred. No. 6.7e-115;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Db 305 VSFGINQAVETNRDNG 320
|||||
RESULT 4
US-08-600-993A-4
; Sequence 4, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-600-993A-4

Query Match 46.5%; Score 1621; DB 2; Length 320;
Best Local Similarity 99.7%; Pred. No. 6.7e-115;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Db 1 KIYDNKNQIADLGSERRVANAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60  
QY LOSNSLOGGSTLTQOLIKLIFYESTSTSDQTSRKAQEAAMLAIQLEQKATKOEILTYINK 120  
Db 61 LOSNSLOGGSTLTQOLIKLIFYESTSTSDQTSRKAQEAAMLAIQLEQKATKOEILTYINK 120  
QY 121 VYNSNGNYGQTAQNTYGGDLNNLSLPOLALLAGMPQAPNOQDPYSHPEAODRRLVL 180  
Db 121 VYNSNGNYGQTAQNTYGGDLNNLSLPOLALLAGMPQAPNOQDPYSHPEAODRRLVL 180  
QY 181 SEMKNQGYISAEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEFGYMLLT 240  
Db 181 SEMKNQGYISAEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEFGYMLLT 240  
QY 241 TGMQVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDSNGKVIQOLGARHOSSN 300  
Db 241 TGMQVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDSNGKVIQOLGARHOSSN 300  
QY 301 VSEGINQAVETNRDMSGTMRPTDYAPALEGYDSTATIVHDEPNYPGTNTPVYMWDR 360  
Db 301 VSEGINQAVETNRDMSGTMRPTDYAPALEGYDSTATIVHDEPNYPGTNTPVYMWDR 360  
QY 361 GYFGNITLQYALQOOSRVAVETLNKGLNRAKTFNLGLGIDYPSIHYSNASISNTTESD 420  
Db 361 GYFGNITLQYALQOOSRVAVETLNKGLNRAKTFNLGLGIDYPSIHYSNASISNTTESD 420  
QY 421 KKYGASSEKMAAAYAAAFANGTYYKPMYIHKVYFSDSEKEFSNVGRAKETAYAMTD 480  
Db 421 KKYGASSEKMAAAYAAAFANGTYYKPMYIHKVYFSDSEKEFSNVGRAKETAYAMTD 480  
QY 481 MKKTVLTYGGRNAYLAMLPOAGKTGTSNTYDEIEHNIKTQFVADDELFACTRYKYSM 540  
Db 481 MKKTVLTYGGRNAYLAMLPOAGKTGTSNTYDEIEHNIKTQFVADDELFACTRYKYSM 540  
QY 541 AVMTGYSNRLLTPLVGNGLTYAAKYRSMNTYLSGSPEDMNIPEGLYRNGEYFKNGAR 600  
Db 541 AVMTGYSNRLLTPLVGNGLTYAAKYRSMNTYLSGSPEDMNIPEGLYRNGEYFKNGAR 600  
QY 601 STWNSPAPQOPPTSTESSSSSDSTSSSTPTSTNNSTTNPNNNTQOQNTTPDOONON 660  
Db 601 STWNSPAPQOPPTSTESSSSSDSTSSSTPTSTNNSTTNPNNNTQOQNTTPDOONON 660  
QY 661 POPAP 666  
Db 661 POPAP 666

RESULT 2  
US-08-481-435-6  
; Sequence 6, Application US/08481435  
; Patent No. 6027906  
; GENERAL INFORMATION:  
; APPLICANT: Balgaresh, Tanjore S  
; APPLICANT: Town, Christine  
; TITLE OF INVENTION: No. 6027906el Polypeptides  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/481,435  
; FILING DATE: 10-JUL-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: IN 580/MAS/94  
FILING DATE: 01-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9404072-2  
FILING DATE: 24-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Steiner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-151  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 682 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-481-435-6

Query Match 99.5%; Score 3466; DB 3; Length 682;  
Best Local Similarity 99.4%; Pred. No. 5,4e-254;  
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRVANAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60  
Db 17 KIYDNKNQIADLGSERRVANAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 76  
QY 61 LOSNSLOGGSTLTQOLIKLIFYESTSTSDQTSRKAQEAAMLAIQLEQKATKOEILTYINK 120  
Db 77 LOSNSLOGGSTLTQOLIKLIFYESTSTSDQTSRKAQEAAMLAIQLEQKATKOEILTYINK 136  
QY 121 VYNSNGNYGQTAQNTYGGDLNNLSLPOLALLAGMPQAPNOQDPYSHPEAODRRLVL 180  
Db 137 VYNSNGNYGQTAQNTYGGDLNNLSLPOLALLAGMPQAPNOQDPYSHPEAODRRLVL 196  
QY 181 SEMKNQGYISAEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEFGYMLLT 240  
Db 197 SEMKNQGYISAEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVEEFGYMLLT 256  
QY 241 TGMQVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDSNGKVIQOLGARHOSSN 300  
Db 257 TGMQVYTNVDOEAKHLMIDYNTDEYVAYPDDELQVASTIVDSNGKVIQOLGARHOSSN 316  
QY 301 VSEGINQAVETNRDMSGTMRPTDYAPALEGYDSTATIVHDEPNYPGTNTPVYMWDR 360  
Db 317 VSEGINQAVETNRDMSGTMRPTDYAPALEGYDSTATIVHDEPNYPGTNTPVYMWDR 376  
QY 361 GYFGNITLQYALQOOSRVAVETLNKGLNRAKTFNLGLGIDYPSIHYSNASISNTTESD 420  
Db 377 GYFGNITLQYALQOOSRVAVETLNKGLNRAKTFNLGLGIDYPSIHYSNASISNTTESD 436  
QY 421 KKYGASSEKMAAAYAAAFANGTYYKPMYIHKVYFSDSEKEFSNVGRAKETAYAMTD 480  
Db 437 KKYGASSEKMAAAYAAAFANGTYYKPMYIHKVYFSDSEKEFSNVGRAKETAYAMTD 496  
QY 481 MKKTVLTYGGRNAYLAMLPOAGKTGTSNTYDEIEHNIKTQFVADDELFACTRYKYSM 540  
Db 497 MKKTVLTYGGRNAYLAMLPOAGKTGTSNTYDEIEHNIKTQFVADDELFACTRYKYSM 556  
QY 541 AVMTGYSNRLLTPLVGNGLTYAAKYRSMNTYLSGSPEDMNIPEGLYRNGEYFKNGAR 600  
Db 557 AVMTGYSNRLLTPLVGNGLTYAAKYRSMNTYLSGSPEDMNIPEGLYRNGEYFKNGAR 616  
QY 601 STWNSPAPQOPPTSTESSSSSDSTSSSTPTSTNNSTTNPNNNTQOQNTTPDOONON 660  
Db 617 STWNSPAPQOPPTSTESSSSSDSTSSSTPTSTNNSTTNPNNNTQOQNTTPDOONON 676  
QY 661 POPAP 666  
Db 677 POPAP 682

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 03:19:04 ; Search time 32.26 Seconds

(without alignments)  
504.261 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KIYDNKNQIADLCSERRVYN.....TOOSNTTDPQONONPQAPQ 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2-6/ptodata/2/1aa/5A-COMB.pep.\*

2: /cgn2-6/ptodata/2/1aa/5B-COMB.pep.\*

3: /cgn2-6/ptodata/2/1aa/6A-COMB.pep.\*

4: /cgn2-6/ptodata/2/1aa/6B-COMB.pep.\*

5: /cgn2-6/ptodata/2/1aa/6C-COMB.pep.\*

6: /cgn2-6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3484	100.0	666	4 US-08-961-083-2	Sequence 2, Appli
2	3466	99.5	682	3 US-08-481-435-6	Sequence 6, Appli
3	1621	46.5	320	2 US-08-245-511-4	Sequence 4, Appli
4	1621	46.5	320	2 US-08-600-993A-4	Sequence 4, Appli
5	636	18.3	731	1 US-08-731-716-2	Sequence 2, Appli
6	567	16.3	828	3 US-08-481-435-2	Sequence 2, Appli
7	555	15.9	850	3 US-08-481-435-10	Sequence 10, Appli
8	526	15.1	823	3 US-08-481-435-4	Sequence 4, Appli
9	514	14.8	844	3 US-08-481-435-7	Sequence 7, Appli
10	512	14.7	844	3 US-08-481-435-8	Sequence 8, Appli
11	489	14.0	835	3 US-08-481-435-9	Sequence 9, Appli
12	406	11.7	77	2 US-08-245-511-24	Sequence 24, Appli
13	406	11.7	77	2 US-08-600-993A-24	Sequence 24, Appli
14	370.5	10.6	532	3 US-08-481-435-12	Sequence 12, Appli
15	370.5	10.6	532	3 US-08-481-435-11	Sequence 11, Appli
16	288.5	8.3	269	2 US-08-771-716-2	Sequence 2, Appli
17	288.5	8.3	269	2 US-09-057-720A-2	Sequence 2, Appli
18	247	7.1	159	3 US-08-481-435-13	Sequence 13, Appli
19	148	4.2	1164	4 US-08-589-756-1	Sequence 1, Appli
20	148	4.2	1164	4 US-09-206-800-1	Sequence 1, Appli
21	146	4.2	1164	4 US-09-206-800-1	Sequence 1, Appli
22	146	4.2	2843	1 US-07-741-940-2	Sequence 2, Appli
23	146	4.2	2843	1 US-08-289-548A-2	Sequence 2, Appli
24	146	4.2	2843	1 US-08-452-654-2	Sequence 2, Appli
25	146	4.2	2843	2 US-08-370-235A-2	Sequence 2, Appli
26	141	4.0	2842	1 US-07-741-940-7	Sequence 7, Appli
27	141	4.0	2842	1 US-08-289-548A-7	Sequence 7, Appli

28	141	4.0	2842	1 US-08-452-654-7	Sequence 7, Appli
29	141	4.0	2843	1 US-08-452-655B-2	Sequence 2, Appli
30	141	4.0	2843	1 US-08-452-655B-7	Sequence 7, Appli
31	141	4.0	2843	3 US-08-450-582-2	Sequence 2, Appli
32	141	4.0	2843	3 US-08-450-582-7	Sequence 7, Appli
33	141	4.0	2973	2 US-08-821-355A-7	Sequence 7, Appli
34	141	4.0	2973	2 US-09-003-687A-7	Sequence 7, Appli
35	141	4.0	2973	2 US-09-136-605-7	Sequence 7, Appli
36	139	4.0	1007	2 US-08-551-459-4	Sequence 4, Appli
37	136.5	3.9	866	2 US-08-483-101-4	Sequence 4, Appli
38	134.5	3.9	610	2 US-08-525-742-8	Sequence 8, Appli
39	133.5	3.8	478	2 US-08-456-670B-40	Sequence 40, Appli
40	133.5	3.8	933	3 US-08-293-728-2	Sequence 2, Appli
41	133.5	3.8	933	3 US-09-421-868-2	Sequence 2, Appli
42	132	3.8	1181	4 US-09-206-898-23	Sequence 23, Appli
43	130.5	3.7	1150	2 US-08-589-756-3	Sequence 3, Appli
44	130.5	3.7	1150	2 US-09-206-800-3	Sequence 3, Appli
45	130.5	3.7	1150	4 US-09-206-898-3	Sequence 3, Appli

## ALIGNMENTS

```

RESULT
1
US-08-961-083-2
: Sequence 2, Application US/08961083
: Patent No. 6159469
:
GENERAL INFORMATION:
: APPLICANT: Chol et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
:
COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
:
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,083
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Strokes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 666 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-08-961-083-2
:
Query Match 100.0%; Score 3484; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.3e-255;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
KIYDNKNQIADLCSERRVNAANDIPDLVAKAISIDHRFFDHRCIDTIRILCAFLRN 60
|||||

```

**THIS PAGE BLANK (USPTO)**

Mon Aug 12 08:40:00 2002

us-08-961-083-2.rag

Page 13

Db 603 ysyfmrnllaiepslkrrkfdvpglrlk--elvdkiptysspsnslp-lpkktds 654

Search completed: August 11, 2002, 04:25:30  
Job time: 4356 sec











```
OY 179 VLSEKNOGYISAEQYERK-----VN--TPITDGLQSLKSASNPAYMDNLYKEVIN 228
XX      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 253 vlylmhkhkrltdkqwedakkidlkanklvnrpeerqndltngdseynsyvfkvkselnn 312
OY 229 Q--VEEETGYNLTITGMQVYTNVDOEAKHLMIDYNTDEYAYPPDELOQVASTYDVNSG 286
DB 313 nkafkdenlagnvlgsklytmdkdvqkflqn--dvngsfyknkdkdqvgaatlildsktg 370
OY 287 KYIAOLGARHOSNSVFCINOAVETNRDMGSTMKPRITDYAPALEGCVYDSTATIVHDEPY 346
DB 371 glvalsggrdfkdvvn--rngatdph-ptgselkflaygpaieamkwtahaidessy 427
OY 347 NYPGNTFVYMWDRGYFSGNITLQYALQOSRNPVAVETLNKY---GLNRKAKTFLNGLGID 402
DB 428 qvddgt--frnydtkshgtvslydalrgsfnpalkwqsvyknagnadapkkfaaklgljn 485
OY 403 YP-SIHYSNAISSNTTESDCKYGASSEKMAAAYAFANGCTYTKMYTHKVFSGSEKE 461
DB 486 yegdlygsevjsgsase-----fslptqlasafaaalangytnahslqkvtrdgetie 539
OY 462 FSNVGTFRAMKETTAYMTDMK-TVLTYGTGRNAYLAMPQAGKTGTSNYTDEIEHNHK 520
DB 540 ydhtchksmsdytlaymleamkylgtfkpyrsayghvsgymmgakqgtqlygaety----- 594
OY 521 TSQEVAPDEL-----FAGYTRKYSMAVMTGYSNRLPLVGNGLTVAAKVRSMNTLYSEG 575
DB 595 -sqynlppnaakdvwvlnqftpqytnsvvmgfsk-----vfkq-----g 631
OY 576 SNP-----EDWNTPGELYNGEFVFNKGRSTWNSPAPQOPSTESSSSSDSTQSST 631
DB 632 enstfghsqgqeypqfilyen--vmkslssr--dgedfkrpssvsgslpslsvsgsqdnt 686
OY 632 TPST---NNSTTTNNNTTQOOSNTPTDOONON 660
DB 687 tnrsthgssdcsansgtagasnmttrsgsqsn 718

RESULT 10
AAV75603
ID AAV75603 standard; Protein: 805 AA.
XX
AC AAV75603;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 791 protein sequence SPQ ID NO:2680.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
PN W09957280-A2.
XX
PD 11-NOV-1999.
XX
PE 30-APR-1999; 99WO-US09346.
XX
PF 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,
Peteresen J, Pizza M, Rapunoli R, Ratti G, Scalato E, Scarselli M;
```

```
PI Tettelin H, Venter JC;
XX
DR WPI: 2000-062150/05.
DR N-PSDB: AA254365.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 2; Page 1270; 1453pp; English.
XX
CC AA253015 to AA254536, AA254577 to AA254615, and AAV74253 to AAV75941
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 805 AA:

Query Match 18.4%; Score 640; DB 21; Length 805;
Best Local Similarity 26.1%; Pred. No. 5.2e-36;
Matches 203; Conservative 121; Mismatches 272; Indels 182; Gaps 21;

OY 2 IYDNKNOLIADGSEERRVNAQANDIPTDLVKAIVSIEDHREFDRGIDPTIRILGAFRLNL 61
DB 59 lysadgevilmgygeqrretkigdfpevlrnavaedkfrfhrhvgdvawgvararaavng 118
OY 62 OSNSLU-GGSTLTQQLIKLTFSTSTSPQTSRKAQEWLAQLQDKATKOEILTYLVYN 120
DB 119 vsqsvgsastltqgvaknfys---sektfrknevllyakiegslskdllelytnq 175
OY 121 VYMSNGNYGMOTAAQNYGKDLNLSLPQALLAQPQAPNPDYPSPEAODRRNLYL 180
DB 176 lylgrraygfasaadqilytnknvrdltlaeaaamlaglpkapsaynpilvpetaklrqkyl 235
OY 181 SEMKNOGYISAEQYERKAVNTPTDGLQSLKSASNPAYMDN-----YLKEYINO-VEE 233
DB 236 nmleekmtlvqgrdqlneel-----hyerfyrkldgsalylvaemvrgelyek 284
OY 234 TGYNLTITGMQVYTNVDOEAK----- 255
DB 285 ygedaytgqfkvytlvradhgvatealrkalnrtfdrgssyrgeaynidlksdveetv 344
OY 256 --HLMIDYNTDEY----- 267
DB 345 sqylsglylvdkmvpavvldvtkkhvvilqlpggrvrlldrralylfaaravnekmgedr 404
OY 268 -----APPDELOQVASTYDVNSNGKYIAOLGARHOSNSVFCINOAVE 310
DB 405 lrrgavirvknmggrwawvavepllgdalsidakgavrtalvggydfniskt---Inrvay 461
OY 311 TNRDMGSTMKPRITDYAPALEGCVYDSTATIVHDEPYNTPGNTN-----TPVYNMDRGYFG 364
DB 462 amrgsgstfkplv-ysaalskqm--tastvndaplsipgkpgnsgvvtcp-knsgrisyg 517
OY 365 NTLQYALQOSRNPVAVETLNKVGINRAKTFPLNGLGIDYPSIHYSNAISSNTTESDCKYRG 424
DB 518 yltirgaltsaknmvslrlmslsgvygaqylrrfgfssseipalslmaliget----- 572
OY 425 ASSEKMAAAYAFANGCTYTKMYTHKVFSDGSEKE-----FSNVGTFRAMKETTAYMM 478
DB 573 -tplkvaeeasvfangryrvshvldklydrdgrlrraqmqplvaegnpqadprrnaylm 631
OY 479 TDMKTVLTYGTGRN-AVYLAAMPQAGKTGTSNYTDEIEHNHITSQEVAPDELFAGYTRK 537
XX
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 04:16:44 ; Search time 214.73 Seconds  
(without alignments)  
1091.692 Million cell updates/sec

Title: US-08-961-083-2  
Perfect score: 3484  
Sequence: 1 KIYDNKNQLIADLGSERRVN.....TQGSNTTPDQONQNPQAPQ 666

Scoring table:  
BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3502263 segs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/2/paa/PCRTUS.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082.COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083.COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084.COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085.COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086.COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087.COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088.COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089.COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090.COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091.COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092.COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093.COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094.COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095.COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096.COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097.COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098.COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099.COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US100.COMB.pep:\*  
25: /cgn2\_6/ptodata/2/paa/US101.COMB.pep:\*  
26: /cgn2\_6/ptodata/2/paa/US60.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3484	100.0	666	19	US-09-536-784-2
2	3484	100.0	666	21	US-09-765-271-2
3	3483	100.0	666	21	US-09-765-272-2
4	3483	100.0	719	15	US-09-583-110-3973
5	3483	100.0	721	19	US-09-107-433-3705
6	3466	99.5	719	1	PCT-US02-03987-13423
7	3466	99.5	719	22	US-09-815-242-13423

8	3466	99.5	719	24	US-10-072-851-13423	Sequence 13423, A
9	1862.5	53.5	823	15	US-09-107-532-5667	Sequence 5667, Ap
10	1862.5	53.5	823	15	US-09-107-532A-5667	Sequence 5667, Ap
11	1751	50.3	778	1	PCT-US02-03987-10728	Sequence 10728, A
12	1751	50.3	778	22	US-09-815-242-10728	Sequence 10728, A
13	1751	50.3	778	24	US-10-072-851-10728	Sequence 10728, A
14	1751	50.3	789	15	US-09-134-000-4939	Sequence 4939, Ap
15	1621	46.5	320	5	US-08-116-541-4	Sequence 4, Appl1
16	1473.5	42.3	771	20	US-09-634-238-229	Sequence 229, App
17	1473.5	42.3	771	20	US-09-634-238-262	Sequence 262, App
18	1361	39.1	462	24	US-10-091-007-92	Sequence 92, Appl
19	817.5	23.5	727	1	PCT-US02-03987-5782	Sequence 5782, Ap
20	817.5	23.5	727	1	PCT-US02-03987-12777	Sequence 12777, A
21	817.5	23.5	727	22	US-09-815-242-5782	Sequence 5782, Ap
22	817.5	23.5	727	22	US-09-815-242-12777	Sequence 12777, A
23	817.5	23.5	727	24	US-10-072-851-5782	Sequence 5782, Ap
24	817.5	23.5	727	24	US-10-072-851-12777	Sequence 12777, A
25	815.5	23.4	730	20	US-09-611-529-4512	Sequence 4512, Ap
26	811.5	23.3	716	26	US-60-253-625-2437	Sequence 2437, Ap
27	811.5	23.3	716	26	US-60-257-931-2821	Sequence 2821, Ap
28	811.5	23.3	716	26	US-60-269-308-3840	Sequence 3840, Ap
29	809.5	23.2	746	15	US-09-134-001C-3214	Sequence 3214, Ap
30	651.5	19.0	675	20	US-09-634-238-228	Sequence 228, App
31	655	18.8	698	21	US-09-739-449-11825	Sequence 11825, A
32	655	18.8	698	22	US-09-803-110-11825	Sequence 11825, A
33	648.5	18.6	814	17	US-09-328-352-4373	Sequence 4373, Ap
34	641	18.4	834	16	US-09-252-991A-32671	Sequence 32671, A
35	636	18.3	730	15	US-09-107-433-4707	Sequence 4707, Ap
36	636	18.3	731	1	PCT-US97-19070-2	Sequence 4707, Ap
37	636	18.3	731	19	US-09-583-110-4720	Sequence 4720, Ap
38	632.5	18.2	855	17	US-09-328-352-6216	Sequence 6216, Ap
39	625	17.9	660	1	PCT-US97-05223-845	Sequence 845, App
40	625	17.9	660	1	PCT-US97-05223-876	Sequence 876, App
41	625	17.9	660	11	US-08-761-318-845	Sequence 845, App
42	625	17.9	660	11	US-08-761-318-876	Sequence 876, App
43	625	17.9	660	12	US-08-824-132-845	Sequence 845, App
44	625	17.9	660	12	US-08-824-132-876	Sequence 876, App
45	625	17.9	660	13	US-08-993-002A-5603	Sequence 5603, Ap

#### ALIGNMENTS

RESULT 1  
US-09-536-784-2  
; Sequence 2, Application US/09536784  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/536,784  
; FILING DATE: 30-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: Oct-30-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Michelle S. Marks  
; REGISTRATION NUMBER: 41,971  
; REFERENCE/DOCKET NUMBER: PB340P3

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 666 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-536-784-2

Query Match 100.0%; Score 3484; DB 19; Length 666;  
Best Local Similarity 100.0%; Pred. No. 3e-289;  
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNLIADLGSERRVNAQANDIPDYLKAIYSIEDHREFDHRGIDTIRILGAFLRN 60  
DB 1 KIYDNKNLIADLGSERRVNAQANDIPDYLKAIYSIEDHREFDHRGIDTIRILGAFLRN 60  
QY 61 LOSNLSOGGSTLTQOLIKLTYFSTSDQTSIRKQAEAMLAIQLEQKATKQEIILTYINK 120  
DB 61 LOSNLSOGGSTLTQOLIKLTYFSTSDQTSIRKQAEAMLAIQLEQKATKQEIILTYINK 120  
QY 121 VYMSNGNGMGTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPEAODRRNLVL 180  
DB 121 VYMSNGNGMGTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPEAODRRNLVL 180  
QY 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLLT 240  
DB 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLLT 240  
QY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYVAPDELOVASTIVDVSNKGVIYAOLGARHSSN 300  
DB 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYVAPDELOVASTIVDVSNKGVIYAOLGARHSSN 300  
QY 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGTNTPYVNMDR 360  
DB 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGTNTPYVNMDR 360  
QY 361 GFYGNITLQYALQOSRNPVAVETLNKVGINRAKTFNLGLGIDYPSIHSVNAISSNTTESD 420  
DB 361 GFYGNITLQYALQOSRNPVAVETLNKVGINRAKTFNLGLGIDYPSIHSVNAISSNTTESD 420  
QY 421 KRYGASSEKMAAAYAAFANGGTYYPMYIHKVYFSDGSEKESNVGTRAMKETTAAYMMTD 480  
DB 421 KRYGASSEKMAAAYAAFANGGTYYPMYIHKVYFSDGSEKESNVGTRAMKETTAAYMMTD 480  
QY 481 MMKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEEIEENHIKTSQFVAPDELFGAYRRKYSM 540  
DB 481 MMKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEEIEENHIKTSQFVAPDELFGAYRRKYSM 540  
QY 541 AVMTGYSNRLLPVLGNGTLVAAKYRSMKTYISEGSNPEDNWIPGLYRNGEFVKNGAR 600  
DB 541 AVMTGYSNRLLPVLGNGTLVAAKYRSMKTYISEGSNPEDNWIPGLYRNGEFVKNGAR 600  
QY 601 STMWSPAPQOPSTSSSSSSSDSTSPSTNNSTTTPNNTNTOOSNTTPPOQONN 660  
DB 601 STMWSPAPQOPSTSSSSSSSDSTSPSTNNSTTTPNNTNTOOSNTTPPOQONN 660  
QY 661 PQPAPQ 666  
DB 661 PQPAPQ 666

RESULT 2  
US-09-765-271-2  
Sequence 2, Application US/09765271  
GENERAL INFORMATION:  
APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,271  
FILING DATE: 22-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/536,784  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/961,083  
FILING DATE: OCT-30-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PB340P3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 666 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-765-271-2

Query Match 100.0%; Score 3484; DB 21; Length 666;  
Best Local Similarity 100.0%; Pred. No. 3e-289;  
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNLIADLGSERRVNAQANDIPDYLKAIYSIEDHREFDHRGIDTIRILGAFLRN 60  
DB 1 KIYDNKNLIADLGSERRVNAQANDIPDYLKAIYSIEDHREFDHRGIDTIRILGAFLRN 60  
QY 61 LOSNLSOGGSTLTQOLIKLTYFSTSDQTSIRKQAEAMLAIQLEQKATKQEIILTYINK 120  
DB 61 LOSNLSOGGSTLTQOLIKLTYFSTSDQTSIRKQAEAMLAIQLEQKATKQEIILTYINK 120  
QY 121 VYMSNGNGMGTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPEAODRRNLVL 180  
DB 121 VYMSNGNGMGTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPEAODRRNLVL 180  
QY 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLLT 240  
DB 181 SEMKNOGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEEETGYNLLT 240  
QY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYVAPDELOVASTIVDVSNKGVIYAOLGARHSSN 300  
DB 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYVAPDELOVASTIVDVSNKGVIYAOLGARHSSN 300  
QY 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGTNTPYVNMDR 360  
DB 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGTNTPYVNMDR 360  
QY 361 GFYGNITLQYALQOSRNPVAVETLNKVGINRAKTFNLGLGIDYPSIHSVNAISSNTTESD 420  
DB 361 GFYGNITLQYALQOSRNPVAVETLNKVGINRAKTFNLGLGIDYPSIHSVNAISSNTTESD 420  
QY 421 KRYGASSEKMAAAYAAFANGGTYYPMYIHKVYFSDGSEKESNVGTRAMKETTAAYMMTD 480

|||||  
Db 421 KKGASSEKMAAAYAFANGGTYYKPMYIHKVYFSDGSEKEFSNVGTRAMKETTAYMTD 480  
QY 481 MKTIVLYTGGRNAYLAWLPQAGKTGTSNTDEIENHIKTSOFVAPDELPAFYTKRYSM 540  
Db 481 MKTIVLYTGGRNAYLAWLPQAGKTGTSNTDEIENHIKTSOFVAPDELPAFYTKRYSM 540  
QY 541 AVWTGYSNRLTPLYGNGLYTAAKAVYRSMTYLSGSGNPEDMNIPEGLYRNGEYFVKNGAR 600  
Db 541 AVWTGYSNRLTPLYGNGLYTAAKAVYRSMTYLSGSGNPEDMNIPEGLYRNGEYFVKNGAR 600  
QY 601 STWNSPAPQOPPESTESSSSSDSSTSSSTPSTNNSTTTNNNTQOONNTTPOOONON 660  
Db 601 STWNSPAPQOPPESTESSSSSDSSTSSSTPSTNNSTTTNNNTQOONNTTPOOONON 660  
QY 661 POPAOP 666  
Db 661 POPAOP 666

## RESULT 3

US-09-765-272-2  
; Sequence 2, Application US/09765272  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/765,272  
; FILING DATE: 22-Jan-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8512  
; TELEFAX: (301) 309-8504  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 666 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-765-272-2

Query Match 100.0%; Score 3484; DB 21; Length 666;  
Best Local Similarity 100.0%; Pred. No. 3e-289;  
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADGSRERRVNAQANDIPTDLVKAIVSIEDHREFDHDGIDTIRILGAFLRN 60  
Db 1 KIYDNKNOLIADGSRERRVNAQANDIPTDLVKAIVSIEDHREFDHDGIDTIRILGAFLRN 60  
QY 61 LOSNSILOGGSTLTQOLIKILTYFSTSDQTSRKAOEAWLAIOLEKATKOEILTYIYINK 120

|||||  
Db 61 LOSNSILOGGSTLTQOLIKILTYFSTSDQTSRKAOEAWLAIOLEKATKOEILTYIYINK 120  
QY 121 VYMSNKNYMGQTAQNAQNYVYGDNLNLSLPOLALLAGPQAPQNVDPYSHPAADRRNLVY 180  
Db 121 VYMSNKNYMGQTAQNAQNYVYGDNLNLSLPOLALLAGPQAPQNVDPYSHPAADRRNLVY 180  
QY 181 SEKNNGYISAEQYERAVNPTITDGLQSLKSASNYAYMDNYLKEVINQVEEETGYNLTL 240  
Db 181 SEKNNGYISAEQYERAVNPTITDGLQSLKSASNYAYMDNYLKEVINQVEEETGYNLTL 240  
QY 241 TGMDEVYTNVDQEAQKHLMDIYNTDEVYAYPPDELOVASTIVDSNGKVIAQLGARRQSSN 300  
Db 241 TGMDEVYTNVDQEAQKHLMDIYNTDEVYAYPPDELOVASTIVDSNGKVIAQLGARRQSSN 300  
QY 301 VSFGINQAVETNNDGKSTAKPTIDYAPALEYGYDSTATIVHDEPNYPTNTNPPVNMDR 360  
Db 301 VSFGINQAVETNNDGKSTAKPTIDYAPALEYGYDSTATIVHDEPNYPTNTNPPVNMDR 360  
QY 361 GYFGNITLQYALQOSRNVPAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSAISSNTTESD 420  
Db 361 GYFGNITLQYALQOSRNVPAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSAISSNTTESD 420  
QY 421 KKGASSEKMAAAYAFANGGTYYKPMYIHKVYFSDGSEKEFSNVGTRAMKETTAYMTD 480  
Db 421 KKGASSEKMAAAYAFANGGTYYKPMYIHKVYFSDGSEKEFSNVGTRAMKETTAYMTD 480  
QY 481 MKTIVLYTGGRNAYLAWLPQAGKTGTSNTDEIENHIKTSOFVAPDELPAFYTKRYSM 540  
Db 481 MKTIVLYTGGRNAYLAWLPQAGKTGTSNTDEIENHIKTSOFVAPDELPAFYTKRYSM 540  
QY 541 AVWTGYSNRLTPLYGNGLYTAAKAVYRSMTYLSGSGNPEDMNIPEGLYRNGEYFVKNGAR 600  
Db 541 AVWTGYSNRLTPLYGNGLYTAAKAVYRSMTYLSGSGNPEDMNIPEGLYRNGEYFVKNGAR 600  
QY 601 STWNSPAPQOPPESTESSSSSDSSTSSSTPSTNNSTTTNNNTQOONNTTPOOONON 660  
Db 601 STWNSPAPQOPPESTESSSSSDSSTSSSTPSTNNSTTTNNNTQOONNTTPOOONON 660  
QY 661 POPAOP 666  
Db 661 POPAOP 666

## RESULT 4

US-09-583-110-3973  
; Sequence 3973, Application US/09583110  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 3973  
; LENGTH: 719  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-3973

Query Match 100.0%; Score 3483; DB 19; Length 719;  
Best Local Similarity 99.8%; Pred. No. 4.2e-289;  
Matches 665; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADGSRERRVNAQANDIPTDLVKAIVSIEDHREFDHDGIDTIRILGAFLRN 60

|||||  
Db 54 KLYDKNKQIADLGSERRNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113  
QY 61 LOSNSLOGGSTLTQOLIKLTFTSTSDQTSRKAQEMALAIQLEOKATKOEILTYINK 120  
Db 114 LOSNSLOGGSTLTQOLIKLTFTSTSDQTSRKAQEMALAIQLEOKATKOEILTYINK 173  
QY 121 VYMSGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVL 180  
Db 174 VYMSGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVL 233  
QY 181 SEMKNQGYISAOYEKAVNTPTDGLQSLKSASNPAYMDNLYKEVINOVEEETGNLLT 240  
Db 234 SEMKNQGYISAOYEKAVNTPTDGLQSLKSASNPAYMDNLYKEVINOVEEETGNLLT 293  
QY 241 TGMADVTTNVDQAKHMDIYNTDEVYAPDDELQVASTIVDVNSGKYIAQLGARHSSN 300  
Db 294 TGMADVTTNVDQAKHMDIYNTDEVYAPDDELQVASTIVDVNSGKYIAQLGARHSSN 353  
QY 301 VSEGINQAVETNRDMSGTMRPTTDYAPALEYGVYDSTATIVHDEPYNPGTNTPYNMNR 360  
Db 354 VSEGINQAVETNRDMSGTMRPTTDYAPALEYGVYDSTATIVHDEPYNPGTNTPYNMNR 413  
QY 361 GFYGNITLQYALQOSRNPAYETLKNVGLNRAKTPNLGLGIDYPSIHNSNAISSNTESD 420  
Db 414 GFYGNITLQYALQOSRNPAYETLKNVGLNRAKTPNLGLGIDYPSIHNSNAISSNTESD 473  
QY 421 KKYGASSEKMAAAYAAAFANGGTYKPMYTHKVFSDGSEKESNVGTRAMKETTAVMMD 480  
Db 474 KKYGASSEKMAAAYAAAFANGGTYKPMYTHKVFSDGSEKESNVGTRAMKETTAVMMD 533  
QY 481 MKKTVLTYGTGNAYLAWLPQAGKTGTSNYTDEIEENHIKTSQFAPAPDELFAGYTRKYSM 540  
Db 534 MKKTVLTYGTGNAYLAWLPQAGKTGTSNYTDEIEENHIKTSQFAPAPDELFAGYTRKYSM 593  
QY 541 AVMTCYSNRLLPLVNGGLTVAAKYVRSMTYLSSEGSNPEDMNIPGLYRNGEFVKNGAR 600  
Db 594 AVMTCYSNRLLPLVNGGLTVAAKYVRSMTYLSSEGSNPEDMNIPGLYRNGEFVKNGAR 653  
QY 601 STWNSPAPQPPSTESSSSSDSSTSPSTNNSTTTPNNTTQOOSNTTTPDOQON 660  
Db 654 STWNSPAPQPPSTESSSSSDSSTSPSTNNSTTTPNNTTQOOSNTTTPDOQON 713  
QY 661 PQAPQ 666  
Db 714 PQAPQ 719

RESULT 5  
US-09-107-433-3705  
: Sequence 3705, Application US/09107433  
: GENERAL INFORMATION:  
: APPLICANT: Lynn A Doucette-Stamm and David Bush  
: TITLE OF INVENTION: NOCLEIC ACID AND AMINO ACID  
: SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE  
: THERAPEUTICS FOR DIAGN  
: NUMBER OF SEQUENCES: 5206  
: CORRESPONDENCE ADDRESS:  
: ADDRESSSEE: GENOME THERAPEUTICS CORPORATION  
: STREET: 100 Beaver Street  
: CITY: Waltham  
: STATE: Massachusetts  
: COUNTRY: USA  
: ZIP: 02354  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: CD-ROM ISO9660  
: COMPUTER: <Unknown>  
: OPERATING SYSTEM: <Unknown>  
: SOFTWARE: <Unknown>  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/107,433  
: FILING DATE: 30-Jun-1998  
: PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/ 085131  
FILING DATE: May 12, 1998  
APPLICATION NUMBER: 60/051553  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arianello, Pamela Deneke  
REGISTRATION NUMBER: 40, 489  
REFERENCE/DOCKET NUMBER: GTC-011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 3705:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 721 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pneumoniae  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...721  
SEQUENCE DESCRIPTION: SEQ ID NO: 3705:  
US-09-107-433-3705

Query Match 100.0%; Score 3483; DB 15; Length 721;  
Best Local Similarity 99.8%; Pred. No. 4.3e-289;  
Matches 665; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KLYDKNKQIADLGSERRNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60  
Db 56 KLYDKNKQIADLGSERRNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 115  
QY 61 LOSNSLOGGSTLTQOLIKLTFTSTSDQTSRKAQEMALAIQLEOKATKOEILTYINK 120  
Db 116 LOSNSLOGGSTLTQOLIKLTFTSTSDQTSRKAQEMALAIQLEOKATKOEILTYINK 175  
QY 121 VYMSGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVL 180  
Db 176 VYMSGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVL 235  
QY 181 SEMKNQGYISAOYEKAVNTPTDGLQSLKSASNPAYMDNLYKEVINOVEEETGNLLT 240  
Db 236 SEMKNQGYISAOYEKAVNTPTDGLQSLKSASNPAYMDNLYKEVINOVEEETGNLLT 295  
QY 241 TGMADVTTNVDQAKHMDIYNTDEVYAPDDELQVASTIVDVNSGKYIAQLGARHSSN 300  
Db 296 TGMADVTTNVDQAKHMDIYNTDEVYAPDDELQVASTIVDVNSGKYIAQLGARHSSN 355  
QY 301 VSEGINQAVETNRDMSGTMRPTTDYAPALEYGVYDSTATIVHDEPYNPGTNTPYNMNR 360  
Db 356 VSEGINQAVETNRDMSGTMRPTTDYAPALEYGVYDSTATIVHDEPYNPGTNTPYNMNR 415  
QY 361 GFYGNITLQYALQOSRNPAYETLKNVGLNRAKTPNLGLGIDYPSIHNSNAISSNTESD 420  
Db 416 GFYGNITLQYALQOSRNPAYETLKNVGLNRAKTPNLGLGIDYPSIHNSNAISSNTESD 475  
QY 421 KKYGASSEKMAAAYAAAFANGGTYKPMYTHKVFSDGSEKESNVGTRAMKETTAVMMD 480  
Db 476 KKYGASSEKMAAAYAAAFANGGTYKPMYTHKVFSDGSEKESNVGTRAMKETTAVMMD 535  
QY 481 MKKTVLTYGTGNAYLAWLPQAGKTGTSNYTDEIEENHIKTSQFAPAPDELFAGYTRKYSM 540  
Db 536 MKKTVLTYGTGNAYLAWLPQAGKTGTSNYTDEIEENHIKTSQFAPAPDELFAGYTRKYSM 595  
QY 541 AVMTCYSNRLLPLVNGGLTVAAKYVRSMTYLSSEGSNPEDMNIPGLYRNGEFVKNGAR 600  
Db 596 AVMTCYSNRLLPLVNGGLTVAAKYVRSMTYLSSEGSNPEDMNIPGLYRNGEFVKNGAR 655  
QY 601 STWNSPAPQPPSTESSSSSDSSTSPSTNNSTTTPNNTTQOOSNTTTPDOQON 660



Db 656 STMNSPAPQPPSTESSSSSDSSTQSSSTPSTNNSTTTNNNTTQOSNTTPOQNON 715  
Qy 661 POPAP 666  
Db 716 POPAP 721

## RESULT 6

PCT-US02-03987-13423  
; Sequence 13423, Application PC/TUS0203987  
; GENERAL INFORMATION:  
; APPLICANT: Elitra Pharmaceuticals, Inc.  
; TITLE OF INVENTION: Methods for identifying the Target of a Compound which Inhibits  
; FILE REFERENCE: ELITRA.028VPC  
; CURRENT APPLICATION NUMBER: PCT/US02/03987  
; CURRENT FILING DATE: 2002-02-02  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 15811  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13423  
; LENGTH: 719  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
PCT-US02-03987-13423

Query Match 99.5%; Score 3466; DB 1; Length 719;  
Best Local Similarity 99.4%; Pred. No. 1.2e-287;

Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFARN 60  
Db 54 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFARN 113  
Qy 61 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEWAMIAIOLEKATQOELITYINK 120  
Db 114 LOSNSLOGGSALTQOLIKLTYFSTSDQTSRKAQEWAMIAIOLEKATQOELITYINK 173  
Qy 121 VYMSNGNYGQTAQONYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPRAODRRNLVL 180  
Db 174 VYMSNGNYGQTAQONYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPRAODRRNLVL 233  
Qy 181 SEKKNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 240  
Db 234 SEKKNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 293  
Qy 241 TGMADVNTNDOEQAKHLMIDYNTDEVYAYPDDELQVASTTVDVSNKGVIQAOLGARHOSN 300  
Db 294 TGMADVNTNDOEQAKHLMIDYNTDEVYAYPDDELQVASTTVDVSNKGVIQAOLGARHOSN 353  
Qy 301 VSEGINQAVETNRDWSGTYMKPTIDYAPALEYGYESTATIVHDEPNYPCTNTPVYNMWR 360  
Db 354 VSEGINQAVETNRDWSGTYMKPTIDYAPALEYGYESTATIVHDEPNYPCTNTPVYNMWR 413  
Qy 361 GYFGNTLTQYALQOOSRNVAVENTLNKVRATFTLNGIDIPSHYSAISSNTTESD 420  
Db 414 GYFGNTLTQYALQOOSRNVAVENTLNKVRATFTLNGIDIPSHYSAISSNTTESD 473  
Qy 421 KYGASSEKMAAAYAFANGGTYRKPMYTHKVVFSGSESEKFSNVGTRAKKETTAAVMTD 480  
Db 474 KYGASSEKMAAAYAFANGGTYRKPMYTHKVVFSGSESEKFSNVGTRAKKETTAAVMTD 533  
Qy 481 MKTIVLTGTGRNAVLAFLPQAGKTGTSNTTDEIENHIKTSQFVAPDELFACTRYKSM 540  
Db 534 MKTIVLTGTGRNAVLAFLPQAGKTGTSNTTDEIENHIKTSQFVAPDELFACTRYKSM 593  
Qy 541 AVWTGNSNLTPLVNGGLVAAKVVSRMNTYLSGNSNPEDMNTPEGLYRNGEEVFENGAR 600  
Db 594 AVWTGNSNLTPLVNGGLVAAKVVSRMNTYLSGNSNPEDMNTPEGLYRNGEEVFENGAR 653  
Qy 601 STMNSPAPQPPSTESSSSSDSSTQSSSTPSTNNSTTTNNNTTQOSNTTPOQNON 660

Db 654 STMNSPAPQPPSTESSSSSDSSTQSSSTPSTNNSTTTNNNTTQOSNTTPOQNON 713  
Qy 661 POPAP 666  
Db 714 POPAP 719

## RESULT 7

US-09-815-242-13423  
; Sequence 13423, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13423  
; LENGTH: 719  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13423

Query Match 99.5%; Score 3466; DB 22; Length 719;  
Best Local Similarity 99.4%; Pred. No. 1.2e-287;  
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFARN 60  
Db 54 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFARN 113  
Qy 61 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEWAMIAIOLEKATQOELITYINK 120  
Db 114 LOSNSLOGGSALTQOLIKLTYFSTSDQTSRKAQEWAMIAIOLEKATQOELITYINK 173  
Qy 121 VYMSNGNYGQTAQONYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPRAODRRNLVL 180  
Db 174 VYMSNGNYGQTAQONYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPRAODRRNLVL 233  
Qy 181 SEKKNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 240  
Db 234 SEKKNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 293  
Qy 241 TGMADVNTNDOEQAKHLMIDYNTDEVYAYPDDELQVASTTVDVSNKGVIQAOLGARHOSN 300  
Db 294 TGMADVNTNDOEQAKHLMIDYNTDEVYAYPDDELQVASTTVDVSNKGVIQAOLGARHOSN 353  
Qy 301 VSEGINQAVETNRDWSGTYMKPTIDYAPALEYGYESTATIVHDEPNYPCTNTPVYNMWR 360

```

Db 354 VSEFGINQAVETNRDMSGSTMKPTITDYAPALEYGYVESTATIVHDEPNYNGTNPVYNNMR 413
Qy 361 GYFGNITTOYALQOOSRNPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNAISSNTESD 420
Db 414 GYFGNITTOYALQOOSRNPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNAISSNTESD 473
Qy 421 KRYGASSEKMAAAYAFANGGTYRPMYTHKVFSDGSEKESNNGTRAMKETTAAYMTD 480
Db 474 KRYGASSEKMAAAYAFANGGTYRPMYTHKVFSDGSEKESNNGTRAMKETTAAYMTD 533
Qy 481 MKKTIVLTGTYGNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFGYTRKYSM 540
Db 534 MKKTIVLTGTYGNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFGYTRKYSM 593
Qy 541 AVMTGYSNRLTPLYNGNGLVAAKYVRSMWTYSEGSNPEDMNIPEGLYRNGEVEKNGAR 600
Db 594 AVMTGYSNRLTPLYNGNGLVAAKYVRSMWTYSEGSNPEDMNIPEGLYRNGEVEKNGAR 653
Qy 601 STWNSPAPQOPPSTESSSSSDSSTSPSTNSSTTTNPNNNTQOQNTTPOQONON 660
Db 654 STWNSPAPQOPPSTESSSSSDSSTSPSTNSSTTTNPNNNTQOQNTTPOQONON 713
Qy 661 POPAOP 666
Db 714 POPAOP 719

RESULT 8
US-10-072-851-13423
: Sequence 13423, Application US/10072851
: GENERAL INFORMATION:
: APPLICANT: Carr, Grant J.
: APPLICANT: Xu, H. Howard
: APPLICANT: Foulkes, J. Gordon
: APPLICANT: Zamudio, Carlos
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Twamick, John D.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Roemer, Terry
: APPLICANT: Jiang, Bo
: APPLICANT: Boone, Charles
: APPLICANT: Bussey, Howard
: TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
: FILE REFERENCE: ELITRA.028A
: CURRENT APPLICATION NUMBER: US/10/072,851
: PRIOR FILING DATE: 2002-02-08
: PRIOR APPLICATION NUMBER: 60/267,636
: NUMBER OF SEQ ID NOS: 15811
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 13423
: LENGTH: 719
: TYPE: PRT
: ORGANISM: Streptococcus pneumoniae
: US-10-072-851-13423

Query Match 99.5%: Score 3466; DB 24: Length 719;
Best Local Similarity 99.4%: Pred. No. 1.2e-287;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```

Qy 121 VYMSNGNVMOTAQONVYCKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVL 180
Db 174 VYMSNGNVMOTAQONVYCKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVL 233
Qy 181 SEMKNQGYISABOYEKAVNPTITDGLQSLKSASNPAYMDNLYKEVINQVEETGYNLLT 240
Db 234 SEMKNQGYISABOYEKAVNPTITDGLQSLKSASNPAYMDNLYKEVINQVEETGYNLLT 293
Qy 241 TGMDEVYTNDOAKHMDIYNTDEVAVYPPDELOVASTIVVSNCKVYLAUGARQSSN 300
Db 294 TGMDEVYTNDOAKHMDIYNTDEVAVYPPDELOVASTIVVSNCKVYLAUGARQSSN 353
Qy 301 VSEFGINQAVETNRDMSGSTMKPTITDYAPALEYGYVESTATIVHDEPNYNGTNPVYNNMR 360
Db 354 VSEFGINQAVETNRDMSGSTMKPTITDYAPALEYGYVESTATIVHDEPNYNGTNPVYNNMR 413
Qy 361 GYFGNITTOYALQOOSRNPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNAISSNTESD 420
Db 414 GYFGNITTOYALQOOSRNPVAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNAISSNTESD 473
Qy 421 KRYGASSEKMAAAYAFANGGTYRPMYTHKVFSDGSEKESNNGTRAMKETTAAYMTD 480
Db 474 KRYGASSEKMAAAYAFANGGTYRPMYTHKVFSDGSEKESNNGTRAMKETTAAYMTD 533
Qy 481 MKKTIVLTGTYGNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFGYTRKYSM 540
Db 534 MKKTIVLTGTYGNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFGYTRKYSM 593
Qy 541 AVMTGYSNRLTPLYNGNGLVAAKYVRSMWTYSEGSNPEDMNIPEGLYRNGEVEKNGAR 600
Db 594 AVMTGYSNRLTPLYNGNGLVAAKYVRSMWTYSEGSNPEDMNIPEGLYRNGEVEKNGAR 653
Qy 601 STWNSPAPQOPPSTESSSSSDSSTSPSTNSSTTTNPNNNTQOQNTTPOQONON 660
Db 654 STWNSPAPQOPPSTESSSSSDSSTSPSTNSSTTTNPNNNTQOQNTTPOQONON 713
Qy 661 POPAOP 666
Db 714 POPAOP 719

RESULT 9
US-09-107-532-5667
: Sequence 5667, Application US/09107532
: GENERAL INFORMATION:
: APPLICANT: Lynn A Doucette-Stamm and David Bush
: TITLE OF INVENTION: NOCLETIC ACID AND AMINO ACID SEQUENCES RELATING TO
: TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR
: TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
: NUMBER OF SEQUENCES: 7308
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02354
: COMPUTER READABLE FORM:
: MEDIUM TYPE: CD-ROM ISO9660
: COMPUTER:
: OPERATING SYSTEM:
: SOFTWARE:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/107,532
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/085598
: FILING DATE: May 14, 1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/051571
: FILING DATE: July 2, 1997
: ATTORNEY/AGENT INFORMATION:
```

NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5667:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 823 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1...823  
US-09-107-532-5667

Query Match 53.5%; Score 1862.5; DB 15; Length 823;  
Best Local Similarity 53.9%; Pred. No. 7e-150;  
Matches 371; Conservative 109; Mismatches 179; Indels 29; Gaps 8;

QY 1 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHRRPDHRCIDTIRILGAFLRN 60  
DB 115 KIYDGDGELFEDLGAERKRISANELPKTLEDIAIVSEDRFRFKHGVDPDIRIGSALS 174  
QY 61 LOSNSLOGSGSTLTQOLIKLYFSTSDOTISRKAOEAMLAIOLEOKATQOELITYYINK 120  
DB 175 FTSGGILOGSGSTLTQOLIKLSFSTSAEDOTLKRKAQAMMAVRLBOKKSKOELITYYINK 234  
QY 121 VYMSNGNGYQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLV 180  
DB 235 VYMSNGLYGEMTASEMYFEKLSLSELPOTALLAGMPQAPSAVDYVYPDOAKRRDYL 294  
QY 181 SEMKNGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVINEVEETGNL 240  
DB 295 YTMLONEKISQREYDQAVNVPYDGLQELTQSDDNKTIYDNYKEVINQOETDKNVT 354  
QY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVAVAPDELAQVASTIVDSNGKVIAGARHOSN 300  
DB 355 DLEIYTNLDLDAOKKLYIVMTDOVSYPDDEMAYASTLIDNTGKVAQIGRIAD 414  
QY 301 VSEGINQAVETNRDMGSTMPTDYAPALEYGVYDSTATIVHDEPNYFGTNTPVYNM 360  
DB 415 YTLGNLAVNTSRDFGSTMKPVTDYGAPEYLYK-STGKTIDAPYNEGSTPVGNM 473  
QY 361 GFNGNTLOAYLAQOSRNVAVELINKVGLNRAKTFPLNGIGIOPSTHYSNATSSNTEED 420  
DB 474 QYMGTTTLOALYLSRNVAVLFNEVGSQVASFKNLGLIESTIHQSNATSSNTEED 533  
QY 421 -KKYASSSEKMAAAYAAAFANGGTYYKPIYHKVFESEKSEPSNGTAMKETTAYMM 479  
DB 534 GKRYGASSLKMAAAAFANGGTYYKPYVKNIVFODGTEERYEPBGKRAMSETYMT 593  
QY 480 DMKRYVLYGTRNAVLAWLPOAGKTGTSNYDEIENHIKTSQVAPDELFAGYTRKYS 539  
DB 594 DILKTLTEGTGTNOAGLYOAGKTGTSNYDDEYAK-LGISSGYVPDILFAGYTPNYS 652  
QY 540 MAVMTGYSRLPIVINGNGLTVAAKYRSMMTYLSSEGSNEDMNIPEGLR-NEEFKFN 597  
DB 653 ISVMYGNKMKMPVSESSHVASDVYRELMOYVANVTMDWEMPSGLIRVGGELYKDO 712  
QY 598 -GARSTWNSPAPQOP-----PSTESSSSSDSTSSQSSSTTSTNNST----- 639  
DB 713 YARSASAIRPTTIPSSSVVQTPGSSSTTETTTQSSSSTQSSSTASSSTTAETSEPA 772  
QY 640 --TTNPNNNTQOQNT---TPDOQONP 661  
DB 773 SSTVPSSSSESSSTPSSSAPASSSEP 800

RESULT 10  
US-09-107-532A-5667  
Sequence 5667, Application US/09107532A  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucelte-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5667:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 823 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...823  
SEQUENCE DESCRIPTION: SEQ ID NO: 5667;  
US-09-107-532A-5667

Query Match 53.5%; Score 1862.5; DB 15; Length 823;  
Best Local Similarity 53.9%; Pred. No. 7e-150;  
Matches 371; Conservative 109; Mismatches 179; Indels 29; Gaps 8;

QY 1 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHRRPDHRCIDTIRILGAFLRN 60  
DB 115 KIYDGDGELFEDLGAERKRISANELPKTLEDIAIVSEDRFRFKHGVDPDIRIGSALS 174  
QY 61 LOSNSLOGSGSTLTQOLIKLYFSTSDOTISRKAOEAMLAIOLEOKATQOELITYYINK 120  
DB 175 FTSGGILOGSGSTLTQOLIKLSFSTSAEDOTLKRKAQAMMAVRLBOKKSKOELITYYINK 234  
QY 121 VYMSNGNGYQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLV 180  
DB 235 VYMSNGLYGEMTASEMYFEKLSLSELPOTALLAGMPQAPSAVDYVYPDOAKRRDYL 294  
QY 181 SEMKNGYISAEQYKAVNTPITDGLQSLKSASNPAYMDNLYKEVINOYEETGNL 240  
DB 295 YTMLONEKISQREYDQAVNVPYDGLQELTQSDDNKTIYDNYKEVINQOETDKNVT 354  
QY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVAVAPDELAQVASTIVDSNGKVIAGARHOSN 300

```

      355 DGLFETYNLMDLDAOKKLYDINTDOYVSPDEMOVASTLIDTNGKKAQIGRHIED 414
      301 VSFQINOVETNRDWMGSMKPTITDAPALLEYGVYSTATIYHDEPYNPGTNTPYNMNR 360
      415 VTLGNLNAVNTSRDPSGSMKPYTDGPAFELYKY-STCKTITDAPYNEGSTPGANDN 473
      361 GYFGNITLQVALOOSRNPVAVETLNKVLNRAKFTPLNGLIDYPSIHSNAISSNTTESD 420
      474 QYMGTITLQALYLSRNPVAVKLENEVSGDKVASFLKNLIGESTIHQSNAISSNTEBED 533
      421 -KKYGASSEKMAAAYAAAFANGCTYKPKYIHKVYFSDSEKEFEFNVGTRAKETAYMMT 479
      534 GTRKGASLLKMAAAYAAAFANGCTYKPKYIHKVYFSDSEKEFEFNVGTRAKETAYMMT 593
      480 DMKFTVLVYGTGRNAYLAMLPOAKGTGTSNTDEIEHNTSQFVAPDELFCAGTTRKYS 539
      594 DILKDTITEGCTGMAQIAGLQOAKGTGTSNTDDEYAK-LGISGVYDILFCAGTTRKYS 652
      540 MAVMTGYSNRLTPLYGNGLYAAKVRSMNTYLSGNSPEDMNIPEGLYR-NGEFVRKN- 597
      653 ISVMTGYDKMKTPTVSESHVADYRELQYVSANVTNTDWMPSGLIRVGGELYKXQ 712
      598 -GARSTWNSPAPQOP-----PSTESSSSSDSSTISQSSSTTPSTNNST----- 639
      713 YTAKSMAITPSTTIPSSSYVQTPGSSSTTETTSQSSSTTSESTAESKESSTTAETSEPA 772
      640 --TTPNNNTQOQNT---TPDOQONP 661
      773 SSTTVPSSSSEESTPSSSAPPASSSEP 800

```

```

RESULT 11
PCT-US02-03987-10728
; Sequence 10728, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; FILE REFERENCE: ELITRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; PRIOR FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10728
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
PCT-US02-03987-10728

```

Query Match 50.38; Score 1751; DB 1; Length 778;  
Best Local Similarity 51.48; Pred. No. 2.4e-140;  
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

```

      1 KIYNNKNOILLADLSEPRVNAQANDIPDLVKAIVSTEDHREFPDHGRGIDTRIIGAFERN 60
      81 KLYDINNEIFEDLGAERKELIQPDVFPQLDAIVSEDRRFYKHIGVDRIITIGSALSN 140
      61 LOSNSLOGGSTLTQOLIKLYFSTSTSDQTSRKAQOEAAMLAIQLEOKATKOELLTYYNK 120
      141 VKNGGLGGSTLTQOLIKLYFSTKESDQTLKRAQEAAMAVRLEREREKREILLTYYNK 200
      121 VYNSNGNYGMOATAQNYGKLNLSLPQLALLAGMPQAPQYDPSHPELAQDRNLVYL 180
      201 VYANGYGYGMEITAENYKHLSELDPQTALLAGMPQAPNSYPYKPDPAKAKERDVL 260
      181 SEMKNQGYISAEQYKAVNPITDGLQSLKGSANVPAMDYULEVINOQVEEETGYNLT 240
      261 YTMIDNKKISAEYKAKATPIDGLVPLKASDNRKRVVDYKAEVINEVAKATGKNYIT 320

```

```

      241 TGMIDYINVDQEAOKHLMIDYNTDEYVYPDELOVASTIVDYSNGKIYQIAGARHOSNN 300
      321 DGLDITNMDMAQKQOLDIVNSDQYVAFPPDKMQVASTIVDVSAGVRAQIGRHIED 380
      301 VSFQINOVETNRDWMGSMKPTITDAPALLEYGVYSTATIYHDEPYNPGTNTPYNMNR 360
      361 VOLGNLNAVNTSRDPSGSMKPYTDGPAFELYKY-STCKTITDAPYNEGSTPGANDN 420
      474 QYMGTITLQALYLSRNPVAVKLENEVSGDKVASFLKNLIGESTIHQSNAISSNTEBED 533
      421 -KKYGASSEKMAAAYAAAFANGCTYKPKYIHKVYFSDSEKEFEFNVGTRAKETAYMMT 479
      534 GTRKGASLLKMAAAYAAAFANGCTYKPKYIHKVYFSDSEKEFEFNVGTRAKETAYMMT 593
      480 DMKFTVLVYGTGRNAYLAMLPOAKGTGTSNTDEIEHNTSQFVAPDELFCAGTTRKYS 539
      594 DILKDTITEGCTGMAQIAGLQOAKGTGTSNTDDEYAK-LGISGVYDILFCAGTTRKYS 652
      540 MAVMTGYSNRLTPLYGNGLYAAKVRSMNTYLSGNSPEDMNIPEGLYR-NGEFVRKN- 597
      653 ISVMTGYDKMKTPTVSESHVADYRELQYVSANVTNTDWMPSGLIRVGGELYKXQ 712
      598 -GARSTWNSPAPQOP-----PSTESSSSSDSSTISQSSSTTPSTNNST----- 639
      713 YTAKSMAITPSTTIPSSSYVQTPGSSSTTETTSQSSSTTSESTAESKESSTTAETSEPA 772
      640 --TTPNNNTQOQNT---TPDOQONP 666
      740 SSAEQPATSEQPEP 754

```

```

RESULT 12
US-09-815-242-10728
; Sequence 10728, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskel, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10728
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10728

```

Query Match 50.3%; Score 1751; DB 22; Length 778;  
Best Local Similarity 51.4%; Pred. No. 2.4e-140;  
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

QY 1 KIYDKKNOLIADLGSERRVNAQANDIPDLVKAIVSIEDHREDFHRGIDTIRILGAFLRN 60  
DB 81 KLYDINNEIFEDLGAERKRELIQPNVDPOLKDAIVSEDRRRYKHIGVDPRIISALSIN 140  
QY 61 LOSNSLOGSSTLTQOLIKLTFFSTSDQTSRKQAQEMALAIQLEQAKTKEILTYIYNK 120  
DB 141 VKNGLGSGSTLTQOLIKLTFFSTSDQTSRKQAQEMALAIQLEQAKTKEILTYIYNK 200  
QY 121 VYMSNGNTGMOTAQNYGKDLNLSLPOLALLAGMPOAPNOYDPSHPEAADRNLVL 180  
DB 201 VYMANGFYEMTAENYGYKHLSELDLPOTALLAGMPOAPNSYDPTKPTAKERRDVL 260  
QY 181 SEMKNOGYISAQYERKAVNPTTDLQSLKSASNPAYMDNLYKEVINOEBETGYNLT 240  
DB 261 YTMVNDKRIKSKAEYKAKATPIDEGLVPLKASDNRKVDNVEVINEVKAKTGNVYT 320  
QY 241 TGMVYTVNDOEAOKHLMIDYNTDEYVAYPPDELOYASTIVDSNGKYIAQLGARHOSN 300  
DB 321 DGLDITYNLDMNAQOKOLYDVNSDQYVAFPPDKMQVASTVIDVASGOVRAQIGRHIPDD 380  
QY 301 VSFGINQAVETNRDMSGTMRKITDYAPALEYGVYDSTATIVHDEPNYNGTNTPYNMNR 360  
DB 381 VOLGNLNAVNTORDVGSYVKPIMDYGPALLENLY-STGRIMVADKPTKPTGIDIDVFNSDL 439  
QY 361 GFYGNITLOYALQOSRNVPAYETLNKVLINRAKTEFLNGLGIDYPSIHNSAITSNTESD 420  
DB 440 TYOGVITMRRAIMGSRNTTAVQTFDEVGKENIMPTKIGIDYKMLEASNAISSNTSDVD 499  
QY 421 -KRYASSEKMAAATAAFANGSTYKPMYIHKVVSDDSEKFEFSNVGRAMKETTAYMYT 479  
DB 500 GDRKYGISLKLAAAYAAAFANNGIYKPYVKNVYFNDGTSVDYQDGRKAMDSTAYMYT 559  
QY 480 DMKATVLTGTRNAYLAWLPQAGKTGTSNNTDEIEHNIKTSQVAPDELFAGTTRKYS 539  
DB 560 DMLKTVLNGTGTFNGAIPGLIOAKTGTSNNTDEDLARMGTKEKJAPDSTVGYTHYA 619  
QY 540 MAVMTGYSNRLPLVGNGLITVAAKYRSMYTYLSEGSNPEDMNIEGLYRNGEYFKNKA 599  
DB 620 VSWMTGYNDRNTPYIOEYGYGLASDVYREIMSTLSQNVSDMDVQDPSVRYVNELEYADA 679  
QY 600 RSTWN-----SPAPOQPSSTESSSSSDSSTSSSTSTSTNNSTTNNNNNTQOOSN 651  
DB 680 YEVPNVQVLPSTSSAPPESSSTVSSSTKEAESSSSSSSSSSAPSSSEAPSTEQPASS 739  
QY 652 TTPDOONONPOPAOP 666  
DB 740 SSAEQPATSEQPPPEP 754

RESULT 13  
US-10-072-851-10728  
; Sequence 10728, Application US/10072851  
; GENERAL INFORMATION:  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Xu, H. Howard  
; APPLICANT: Foulkes, J. Gordon  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Tirawick, John D.  
; APPLICANT: Roemero, Robert T.  
; APPLICANT: Roemer, Terry  
; APPLICANT: Jlang, Bo  
; APPLICANT: Boone, Charles  
; APPLICANT: Bussey, Howard  
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C  
; TITLE OF INVENTION: Proliferation

FILE REFERENCE: ELITRA\_028A  
; CURRENT APPLICATION NUMBER: US/10/072,851  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 15811  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10728  
; LENGTH: 778  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-10-072-851-10728

Query Match 50.3%; Score 1751; DB 24; Length 778;  
Best Local Similarity 51.4%; Pred. No. 2.4e-140;  
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

QY 1 KIYDKKNOLIADLGSERRVNAQANDIPDLVKAIVSIEDHREDFHRGIDTIRILGAFLRN 60  
DB 81 KLYDINNEIFEDLGAERKRELIQPNVDPOLKDAIVSEDRRRYKHIGVDPRIISALSIN 140  
QY 61 LOSNSLOGSSTLTQOLIKLTFFSTSDQTSRKQAQEMALAIQLEQAKTKEILTYIYNK 120  
DB 141 VKNGLGSGSTLTQOLIKLTFFSTSDQTSRKQAQEMALAIQLEQAKTKEILTYIYNK 200  
QY 121 VYMSNGNTGMOTAQNYGKDLNLSLPOLALLAGMPOAPNOYDPSHPEAADRNLVL 180  
DB 201 VYMANGFYEMTAENYGYKHLSELDLPOTALLAGMPOAPNSYDPTKPTAKERRDVL 260  
QY 181 SEMKNOGYISAQYERKAVNPTTDLQSLKSASNPAYMDNLYKEVINOEBETGYNLT 240  
DB 261 YTMVNDKRIKSKAEYKAKATPIDEGLVPLKASDNRKVDNVEVINEVKAKTGNVYT 320  
QY 241 TGMVYTVNDOEAOKHLMIDYNTDEYVAYPPDELOYASTIVDSNGKYIAQLGARHOSN 300  
DB 321 DGLDITYNLDMNAQOKOLYDVNSDQYVAFPPDKMQVASTVIDVASGOVRAQIGRHIPDD 380  
QY 301 VSFGINQAVETNRDMSGTMRKITDYAPALEYGVYDSTATIVHDEPNYNGTNTPYNMNR 360  
DB 381 VOLGNLNAVNTORDVGSYVKPIMDYGPALLENLY-STGRIMVADKPTKPTGIDIDVFNSDL 439  
QY 361 GFYGNITLOYALQOSRNVPAYETLNKVLINRAKTEFLNGLGIDYPSIHNSAITSNTESD 420  
DB 440 TYOGVITMRRAIMGSRNTTAVQTFDEVGKENIMPTKIGIDYKMLEASNAISSNTSDVD 499  
QY 421 -KRYASSEKMAAATAAFANGSTYKPMYIHKVVSDDSEKFEFSNVGRAMKETTAYMYT 479  
DB 500 GDRKYGISLKLAAAYAAAFANNGIYKPYVKNVYFNDGTSVDYQDGRKAMDSTAYMYT 559  
QY 480 DMKATVLTGTRNAYLAWLPQAGKTGTSNNTDEIEHNIKTSQVAPDELFAGTTRKYS 539  
DB 560 DMLKTVLNGTGTFNGAIPGLIOAKTGTSNNTDEDLARMGTKEKJAPDSTVGYTHYA 619  
QY 540 MAVMTGYSNRLPLVGNGLITVAAKYRSMYTYLSEGSNPEDMNIEGLYRNGEYFKNKA 599  
DB 620 VSWMTGYNDRNTPYIOEYGYGLASDVYREIMSTLSQNVSDMDVQDPSVRYVNELEYADA 679  
QY 600 RSTWN-----SPAPOQPSSTESSSSSDSSTSSSTSTSTNNSTTNNNNNTQOOSN 651  
DB 680 YEVPNVQVLPSTSSAPPESSSTVSSSTKEAESSSSSSSSAPSSSEAPSTEQPASS 739  
QY 652 TTPDOONONPOPAOP 666  
DB 740 SSAEQPATSEQPPPEP 754

RESULT 14  
US-09-134-000-4939  
; Sequence 4939, Application US/09134000A  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS

;; TITLE OF INVENTION: FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
;; FILE REFERENCE: GTC-005  
;; CURRENT APPLICATION NUMBER: US/09/134.000A  
;; CURRENT FILING DATE: 1998-08-13  
;; NUMBER OF SEQ ID NOS: 6810  
;; SEQ ID NO 4939  
;; LENGTH: 789  
;; TYPE: PR1  
;; ORGANISM: Enterococcus faecalis  
US-09-134-000-4939

Query Match 50.3%; Score 1751; DB 15; Length 789;  
Best Local Similarity 51.4%; Pred. No. 2.4e-140;  
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

QY 1 KIYDNKNLIDLIGSERVNAQANDIPTDLVKAIVSIEDHFRFDRGIDTRIIGAFLRN 60  
DB 92 KLYDNNIEFDLGEKRELIOQNDVPOLDALVVEDRRYKHIGVDPRIIGALSALN 151  
QY 61 LOSNSLOGSGLTQOLIKLTYFSTSDQITSRKQAEAMLAIOLEOKATKOEILTYINK 120  
DB 152 VKNGLOGSGLTQOLIKLSTFKESQDITLKRQAEAMAVRLERESKEIILTYINK 211  
QY 121 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPAPNOYDPSHPPEAAODRRNLVL 180  
DB 212 VYMANGFYGMETAENYGYKHLSELDELPTQALLAGMPAPNSYDPTKPTAKERRDYVL 271  
QY 181 SEMKNOGITSADQEKAVNTPITDGLSLKSASNPAYMDNYLKEVINQVEETGYNLLT 240  
DB 272 YTMVDMNKISAEYKAEKATPIDGLVPLKASDDRKVYDVYKVEIVAKTKGNVYT 331  
QY 241 TGMDEVTVNDDQAKHLMIDYNTDEYVAYPPDELOVASTIVDVNSGKVIYLAOLGARHOSN 300  
DB 332 DGLDIYTNVDDQAKHLMIDYNTDEYVAYPPDELOVASTIVDVNSGKVIYLAOLGARHOSN 391  
QY 301 VSEGINQAVETNRDMGSTMPTTDYAPALEXYVDSTAFIVHDEBYNPGNTPYVYNDNR 360  
DB 392 VOLNNNLAVNTOBQVGVSTVFKPIMDYGPALINDLY-STGLRLAWDKTKYKPTGTDIDVFNSDL 450  
QY 361 GYFENITLOVLAQOSRNPAYETLNVKLNRAKTFNLGLIDGIDPSIHSNMISSSTTSD 420  
DB 451 TYQGVITMRRAIMGSRNTTAQOTFDEVGENIMPIKGLGIDYKNLEASNMISSSTTSD 510  
QY 421 -KRYGASSEKMAAAYAAAFANGCTYKPMYIHKVYVSDSESEKFSNVGRAKETAYAMT 479  
DB 511 GDXGIGISLKLAAAYAAAFANGCTYKPMYIHKVYVSDSESEKFSNVGRAKETAYAMT 570  
QY 480 DMKTVLTYIGRNAYLAMLPOAGKTGTSNTDEIENHITSQFVAPDELPAGYTRKYS 539  
DB 571 DMLDVLNGGTGFNGALPGLIOAAKTGTSNTDEIENHITSQFVAPDELPAGYTRKYS 630  
QY 540 MAVVTGYSNRLTPLYGNGLTVAARVYRSMNTYLSSESNPEDMNTPEGLYRNGEYFNKGA 599  
DB 631 VSVATGINDRNTPLYOEYTGASDVIREIMSYLSOONVSNDDWQDVSVYRNGELIYVADA 690  
QY 600 RSTWN-----SPAQPPSTESSSSSDSTSSQSSSTPTNNSTPTNNNNTOOSN 651  
DB 691 YEYVNVQVLPSTSSAPQESSSTYESSSTKEAESSSSSSSESAPSSSEAPRSTEDPASS 750  
QY 652 TTPDOQNONPOPAOP 666  
DB 751 SSAPQATSEOPPEP 765

RESULT 15  
US-08-116-541-4  
; Sequence 4, Application US/08116541  
; GENERAL INFORMATION:  
; APPLICANT: Masure, H. Robert  
; APPLICANT: Pearce, Barbara J.  
; APPLICANT: Toumanen, Elaine  
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND

;; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON  
;; NUMBER OF SEQUENCES: 45  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Klauber & Jackson  
;; STREET: 411 Hackensack Avenue  
;; CITY: Hackensack  
;; STATE: New Jersey  
;; COUNTRY: USA  
;; ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/116,541  
FILING DATE: 19930901  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-069  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 320 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-116-541-4

Query Match 46.5%; Score 1621; DB 5; Length 320;  
Best Local Similarity 99.7%; Pred. No. 7.2e-130;  
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIYDNKNLIDLIGSERVNAQANDIPTDLVKAIVSIEDHFRFDRGIDTRIIGAFLRN 60  
DB 5 KIYDNKNLIDLIGSERVNAQANDIPTDLVKAIVSIEDHFRFDRGIDTRIIGAFLRN 64  
QY 61 LOSNSLOGSGLTQOLIKLTYFSTSDQITSRKQAEAMLAIOLEOKATKOEILTYINK 120  
DB 65 LOSNSLOGSGLTQOLIKLTYFSTSDQITSRKQAEAMLAIOLEOKATKOEILTYINK 124  
QY 121 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPAPNOYDPSHPPEAAODRRNLVL 180  
DB 125 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPAPNOYDPSHPPEAAODRRNLVL 184  
QY 181 SEMKNOGITSADQEKAVNTPITDGLSLKSASNPAYMDNYLKEVINQVEETGYNLLT 240  
DB 185 SEMKNOGITSADQEKAVNTPITDGLSLKSASNPAYMDNYLKEVINQVEETGYNLLT 244  
QY 241 TGMDEVTVNDDQAKHLMIDYNTDEYVAYPPDELOVASTIVDVNSGKVIYLAOLGARHOSN 300  
DB 245 TGMDEVTVNDDQAKHLMIDYNTDEYVAYPPDELOVASTIVDVNSGKVIYLAOLGARHOSN 304  
QY 301 VSEGINQAVETNRDMG 316  
DB 305 VSEGINQAVETNRDMG 320

Search completed: August 11, 2002, 04:30:05  
Job time: 801 sec

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 04:23:44 ; Search time 142.66 Seconds  
(without alignments)  
818.374 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KLYDKKNLIDLGSERRVN.....TQGSMTTPDQGNQNPAP 666

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 528882 seqs, 175299045 residues

Total number of hits satisfying chosen parameters: 528882

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: Pending Patents AA New:  
2: /cgn2\_6/ptodata/2/paa/PCP\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US05\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3484	100.0	666	4 US-08-961-083-2	Sequence 2, Appli
2	529.5	15.2	821	5 US-09-769-744A-122	Sequence 122, App
3	526	15.1	789	5 US-09-540-209B-8941	Sequence 8941, Ap
4	384	11.0	790	5 US-09-540-209B-9358	Sequence 9358, Ap
5	162	4.6	875	5 US-09-914-543-6	Sequence 6, Appli
6	161	4.6	738	6 US-10-179-131-10183	Sequence 10183, A
7	160.5	4.6	2893	5 US-09-882-227-522	Sequence 522, App
8	159	4.6	615	5 US-09-882-327-408	Sequence 408, App
9	153.5	4.4	930	5 US-09-200-650D-3	Sequence 3, Appli
10	152	4.4	973	6 US-10-179-131-5286	Sequence 5286, Ap
11	148.5	4.3	2283	1 PCT-US02-19220-4	Sequence 4, Appli
12	148	4.2	188	5 US-09-540-209B-6592	Sequence 6592, Ap
13	147.5	4.2	595	5 US-09-673-605A-34	Sequence 34, Appli
14	146	4.2	2843	5 US-09-442-489B-2	Sequence 2, Appli
15	145.5	4.2	1028	6 US-10-179-131-6430	Sequence 6430, Ap
16	143	4.1	1010	5 US-09-914-543-10	Sequence 10, Appli
17	141	4.0	2842	5 US-09-442-489B-7	Sequence 7, Appli
18	141	4.0	2843	4 US-08-681-219A-30	Sequence 30, Appli
19	141	4.0	2843	5 US-09-987-482-1	Sequence 1, Appli
20	141	4.0	2843	7 US-60-389-987-1970	Sequence 1970, Ap
21	139.5	4.0	706	7 US-60-360-039-1865	Sequence 1865, Ap
22	137	3.9	2297	7 US-60-385-568-217	Sequence 217, App
23	136.5	3.9	567	7 US-60-385-568-226	Sequence 226, App
24	136.5	3.9	567	6 US-10-179-131-7530	Sequence 7530, Ap
25	134	3.8	994	5 US-09-935-625-7547	Sequence 7547, Ap
26	134	3.8	994	5 US-09-935-625-28205	Sequence 28205, A

27	134	3.8	1006	5 US-09-935-625-7546	Sequence 7546, Ap
28	134	3.8	1006	5 US-09-935-625-28204	Sequence 28204, A
29	133.5	3.8	478	5 US-09-445-289B-34	Sequence 34, Appli
30	133.5	3.8	541	7 US-60-360-039-9976	Sequence 9976, Ap
31	133.5	3.8	627	7 US-60-360-039-18499	Sequence 18499, A
32	132	3.8	1014	7 US-60-360-039-1608	Sequence 1608, Ap
33	131.5	3.8	1084	7 US-60-360-039-11155	Sequence 11155, A
34	131.5	3.8	1518	7 US-60-360-039-22243	Sequence 22243, A
35	131.5	3.8	10431	1 PCT-US02-11734-310	Sequence 310, App
36	131	3.8	989	5 US-09-935-625-7548	Sequence 7548, Ap
37	131	3.8	989	5 US-09-935-625-28206	Sequence 28206, Ap
38	131	3.8	1031	5 US-09-935-625-8784	Sequence 8784, Ap
39	131	3.8	1031	5 US-09-935-625-29529	Sequence 29529, A
40	131	3.8	1043	5 US-09-935-625-8783	Sequence 8783, Ap
41	131	3.8	1043	5 US-09-935-625-29528	Sequence 29528, A
42	131	3.8	1332	6 US-10-179-131-9843	Sequence 9843, Ap
43	129	3.7	708	5 US-09-540-209B-6575	Sequence 6575, Ap
44	128	3.7	1029	5 US-09-540-209B-10169	Sequence 10169, Ap
45	127.5	3.7	623	6 US-10-179-131-5358	Sequence 5358, Ap

#### ALIGNMENTS

RESULT 1  
US-08-961-083-2  
Sequence 2, Application US/08961083

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961.083

FILING DATE: 30-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36, 373

REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 666 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-08-961-083-2

Query Match 100.0%; Score 3484; DB 4; Length 666;  
Best Local Similarity 100.0%; Pred. No. 8e-225;  
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KLYDKKNLIDLGSERRVNAQNDIPDVKAIYSIDHREFDHRGIDTIRILGAPLRN 60

```
|||||
Db 1 KIDNKKQQLIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFPDRGIDTIRILGAFLRN 60
QY LOSNSLOGSSTLTFQOLKILYFSTSTSDOTISRAQOEAHLAIOLEOKATQOELLTYINK 120
Db 61 LOSNSLOGSSTLTFQOLKILYFSTSTSDOTISRAQOEAHLAIOLEOKATQOELLTYINK 120
QY 121 VYNSNGNYGQTAQNYVYGRDLNLSLPOLALLAGMPQAPNQYDPSHPPEAODRNLVL 180
Db 121 VYNSNGNYGQTAQNYVYGRDLNLSLPOLALLAGMPQAPNQYDPSHPPEAODRNLVL 180
QY 181 SEKNNGYISAEQYERKAVNPITDGLSLKSASNPAYMDNYLKEVINQVEEFGYMLT 240
Db 181 SEKNNGYISAEQYERKAVNPITDGLSLKSASNPAYMDNYLKEVINQVEEFGYMLT 240
QY 241 TGMADVNTNDOEAKHLMIDYNTDEYVAYPPDELOVASTIVDSNGVIAQLGARHOSN 300
Db 241 TGMADVNTNDOEAKHLMIDYNTDEYVAYPPDELOVASTIVDSNGVIAQLGARHOSN 300
QY 301 VSEGINQAVETNRDMSGTMKPTDYAPALEGYVYDSTATIVHDEPNYPGTNPFVYNMDR 360
Db 301 VSEGINQAVETNRDMSGTMKPTDYAPALEGYVYDSTATIVHDEPNYPGTNPFVYNMDR 360
QY 361 GYRGNTILOALQOOSRNPVPEVLNKGILNRAKTFNLGLGIDPFSIHYNAISSNTTESD 420
Db 361 GYRGNTILOALQOOSRNPVPEVLNKGILNRAKTFNLGLGIDPFSIHYNAISSNTTESD 420
QY 421 KKGASSEKMAAAYAAAFANGGTYKKPMYIHKVYVSDGSEKEFSVNGRAKETTAYAMTD 480
Db 421 KKGASSEKMAAAYAAAFANGGTYKKPMYIHKVYVSDGSEKEFSVNGRAKETTAYAMTD 480
QY 481 MKKTVLTYGTGRNAYLAMLPOAGKTGTSNTYDEIEHNIKTQFVAPDELPAQYTRKYSM 540
Db 481 MKKTVLTYGTGRNAYLAMLPOAGKTGTSNTYDEIEHNIKTQFVAPDELPAQYTRKYSM 540
QY 541 AVMTGTSNRLTPLYGNCILTYAAKYRSMATYLLSEGSNPEDEMNIEGILYRNGEYFKNGAR 600
Db 541 AVMTGTSNRLTPLYGNCILTYAAKYRSMATYLLSEGSNPEDEMNIEGILYRNGEYFKNGAR 600
QY 601 STYNSPAPQOPSTESSSSSDSTSSSTPSTNNSTTNPNNNQOANTPPDOONON 660
Db 601 STYNSPAPQOPSTESSSSSDSTSSSTPSTNNSTTNPNNNQOANTPPDOONON 660
QY 661 POPAOP 666
Db 661 POPAOP 666

RESULT 2
US-09-769-744A-122
; Sequence 122, Application US/09769744A
; GENERAL INFORMATION:
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hamill, Sean B
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PMC/P21122WO
; CURRENT APPLICATION NUMBER: US/09/769,744A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 9816336.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: US 60/125329
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 122
; LENGTH: 821
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-122
```

```
Query Match 15.2%; Score 529.5; DB 5; Length 821;
Best Local Similarity 27.2%; Pred. No. 4.4e-27;
Matches 203; Conservative 115; Mismatches 270; Indels 159; Gaps 32;

QY 9 LIADLGS-RRVNAQANDIPTDLVKAIVSIEDHRRFPDRGIDTIRILGAFLRN 63
Db 115 VIASIESDLRTSISSHQISENLKKAIIATEDEHEKKGVPKAVIRATLGKRV-GLGS 173
QY 64 NSLOGSSTLTFQOLKILYFSTSTSDOTISRAQOEAHLAIOLEOKATQOELLTYINKVYM 123
Db 174 SS--GGSTLTFQOLKILYFSTSTSDOTISRAQOEAHLAIOLEOKATQOELLTYINKVYM 123
QY 124 SNGNYGM-----QTAQNYGQKDLNLSLPOLALLAGMPQAPNQYDPSHP----- 169
Db 229 GRNKGQINAGARQASGICGVASOLTVQOALAGLPOSPTISYENTGELKSDEDL 288
QY 170 EAAQDRNLVLEKKNQGYISAEQYERKAVNPIT-----TDGLQSL 209
Db 289 EIGLRRAKAVLYSMYRTGALSDEYSGYKDYDLKQDFLPSGTVTGISRDYLYFTTLAEQ 348
QY 210 KSASNPAYMDNYLKEVINQVE-----EETGNLITGMDVITNVDQAKH 256
Db 349 ERMVDYLAORDNYSAKELKNEATQKRYDLAKEIENGKYKITTT-----IDQKTHSA 401
QY 257 LMDIYNTDEYVAYPPD---ELOVASTIVDSNGVIAQLGARHOSNVSEGINQAVETNR 313
Db 402 MQSA--VADYGYLLDDTGRVNGVNLMDNOTGAILPFVGRNIOEQNON--NHAFTKR 456
QY 314 DWGSTMKPTDYAPALEGYVYDSTATIVHDEPNYPGTNPFVYNMDRGFGNTILOALQ 373
Db 457 SPASTKPLAYGIAIDQGLMS--ETILSNYPFANGNPIMVANSKG-TGMWTLGBALN 514
QY 374 QSRNPVPEVLNKGILNRAKTFNLGLGIDPFSIHYNAISSNTTESDKKGASSEKM 430
Db 515 YSMNIPAWMTYRMIREKGVV-VKGYMEKMEYEIP-----EYGIELPM 556
QY 431 AA-----AYAAFANGGTYKKPMYIHKVYVSDGS-EKESNVGTRAMKETTYAMM 478
Db 557 GGGIEYVVAOHTNGYQYLANGVYHOKHVISKIEADGRVYVEYQDPVQYISKATITIM 616
QY 479 TDMKTVLTYG-----TGRNAYLAMLPOAGKTGTSNTYDEIEHNIKTQFVAPDE 529
Db 617 OGILREVLSRVTTTFKSNLTSINPLANADWIGKTGTN-QDENMMLMLSTPLTLG- 674
QY 530 LFAQYTRKYSMAVWTGYSN-----RLTPLY-GN-GLTVAAKYRSMATYLLSE 574
Db 675 -WIGHDHNHLSRRAGYSNNSNYMAHLVNAIQASPISIMGNERPALDPSVYKSEV-LKST 732
QY 575 GSNPEDEMNIP-----EGLYRNGEYFKNGARST-----WNSPAPQOP- 611
Db 733 GQKPEKYSVGEKEVEYGVTSYMWANKSGAPATISYFALIGSDADYQNMASSTIVGSLPT 792
QY 612 PSTESSSSSDSTSSSTPSTNNSTTNPNNNQOANTPPDOONON 660
Db 793 PSSSSSSSSSDSTSSSTPSTNNSTTNPNNNQOANTPPDOONON 660

RESULT 3
US-09-540-209B-8941
; Sequence 8941, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 8941
; LENGTH: 789
; TYPE: PRF
```

ORGANISM: B. fragilis  
US-09-540-209B-8941

Query Match 15.1%; Score 526; DB 5; Length 789;  
Best Local Similarity 24.7%; Pred. No. 7.1e-27;  
Matches 177; Conservative 118; Mismatches 248; Indels 174; Gaps 25;

```
QY 13 LGSERVNAQANDIPDYLKATVISEDHREFDHRGIDTIRILGAFIRN--LGSNSLQGS 70
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 77 LSKENNVYTSYNELSPNINALLATEDVRETHSGIDAKALLRAVVRGLLMKNAGGS 136
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 71 TLTQOIKLTYESTSTSDOTISRKAOE--AW-LAIQEQATKOEILTYINKVYNSGN 127
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 137 TISQOLAK-QLTDEVARNTLQRLPQKPIEMVIAVKLEKYKKEILLSMTLKNKFDL 195
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 128 YGMQTAONYGKDLNNLSLPOLLALLAGMPQAPNOYDPYSHPEAADRNLVLEKNKG 187
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 196 VGIKTASYTYFGCEPKDLIEQAATLIGMCKNPSLYNPFNERSHGRRTVLDQRRKAG 255
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 188 YISAEQEKAVNPI-----TDGL-----QSLKASNTPRAY-MDNLYK 224
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 256 YITAEBCDSLQNLPLELVHRYDHKEGLATYEREVLRYGVATSKPVRSNRGMQKFE 315
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 225 EYI-----NOVEETGYMLTTGMDVYTNVDQAKHLMDI----- 260
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 316 DSDIDEMNPLVGMCEKKNKKGDSNVIYTDGLKIYTTINSHQRYADEAVEEHVGELOP 375
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 261 -----YNTDEVYAVPDELO----- 275
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 376 LFEKERRKKAPYSNQLTOEIDRLIDRAVKQTSRYQMKKEAGISEAEIKAFNKPESY 435
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 276 -----VASTI-----VDVNGVYINQAGRHOSNSPFGINQA 308
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 436 SVFTWHGVADFTTSPDSTRYKHKPLRAGFMSMDPIINGOKAVAGGPNVT--YFOYDMA 492
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 309 VETNRDMSGTMKPTDYAPALEYGV--YDST--ATIYHDEPYNPGNTPYVNDRCY 362
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 493 MGRGRQVSGTIRPYL-YALAMENGFSPOCETRNVETLLIDENCKPKSPKNT-----SKGH 546
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 363 FGN-ITLOYALQOOSRNPVAVETLNKYNLRAKTFELNGLGDIDPSIHYSNASSNTTESDK 421
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 547 YGEMVTLKGLANSNMWISAYILMSKINPYALARLISHFGVRNKEIQPTVSLCIGPCE--- 603
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 422 KYGASSEKMAAAYAFANGTYYKPMYIHKVYFSDGSE-KEPSNVTBAMKETTAAMD 480
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 604 ---ISVGENVSATYAFANGIRVAPLEVTKIEDSEGNVLATFSPQMEEVISASAYKMLV 660
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 481 MKKTVLTGTGRNA--YLAMLPQAGKTGTSNTDEIEENHIKTSQFVAPDELFGYTRKY 538
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 661 MLRAVINETGCAVRKRYGTADMGKGTGTTNRN-----DGMFGFTPLSL 705
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 539 SMAVMTGYSNR---LTPLVGNGLVAAVYRSNM--TYLSE--GSNP-EDMNIPBG 586
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 706 VSGCWVGGEERDHDFTMTYGGGASLALPIWTKYMHKVVADQLGYDPKETFNLPGD 762
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
```

## RESULT 4

US-09-540-209B-9358  
; Sequence 9358, Application US/09540209B  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR  
; FILE REFERENCE: 2709.1001-001  
; CURRENT APPLICATION NUMBER: US/09/540,209B  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 10444  
; SEQ ID NO 9358  
; LENGTH: 790  
; TYPE: PRN  
; ORGANISM: B. fragilis  
US-09-540-209B-9358

Query Match 11.0%; Score 384; DB 5; Length 790;  
Best Local Similarity 24.0%; Pred. No. 2.2e-17;  
Matches 136; Conservative 102; Mismatches 240; Indels 88; Gaps 21;

```
QY 2 IYNNKQLIADLSESRVNAQ-----ANDIPTLVAAIVSIEDHREFDHRGIDTIRILGA 56
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 66 VYDRNEL---LGARLATDQGMFFPRENPEVAVCLIEFERORYYHHMGVAPLAIGRA 122
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 57 FLRNLOSNSL-QGSPYLQOOLIKLTYFSTSDOTISRKAQEAMLAIQEQATKOEIL 115
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 123 VYONLKHKRIVSGSTLTMTQIRL---ARNKPRTEKELIEMVATRLFEFRSKKEILS 178
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 116 YITNKYMSNGNMGMTAONTYGKDLNNLSLPOLLALLAGMPAPQYDPYSHPEAADR 175
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 179 LYISHAPFGGNVGLDAAMRYFGHSAEELSMASAMLAVALPNSPAMHLSKRALDLK 238
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 176 RNLVLESMKNQGYISAEQEKAVNPIYTDGLQSLKASNPAYMDVYLKEVINQVEETG 235
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 239 RNLVLHLHKKGLIDTSTYELAISEPLP---QEPPLPPIHAPILTDYFYOTRN-----GK 290
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 236 YNLTTGMDVYTNVDQAKHLMDIYNTDEVYAVPDELOVASTIVDSNGKVIQALGAR 295
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 291 YSVSTIDRGIQTOIESLVER--WN---SEFKR--SDIRRLALVIDIRTNQAIAYCGNV 342
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 296 HOSSNVSFGINQAVETNRDMSGTMKPTDYAPALEYGVYDSTIYHDEPYNPGNTPY 355
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 343 HFDKESQGNQVDVIRSPRSTGSLKPLLYAMQEEELIPN--TLPLDIPVNIENG-TP- 398
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 356 YNMDRGYFNTILOYALQOOSRNPVAVETLNKYNLRAKTFELNGLQ---IDPSIHSNAT 412
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 399 QNFNLQPEGAVRSEALASLNPSTYMLQRYGVPFHSLKIGLTLNRPSSHGLSL 458
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 413 SSNTTESDKKYGASSEKMAAAYAFANGTYYK--PMY-----IHKVYFSDGSE 459
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 459 I-----LGAEAETLMDITSAYANMGRSLNRLPOFCTLLDLSISVHRSPSGAV 509
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 460 KESNVTBAMKETTYAMTDMMKYLYTGTGNALAMLPQAGKTGTSNTDEIEENHI 519
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 510 WQTFD---AIKVNRPREDID-WRTIPSMQT-----IAW-----KGTGS----- 543
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
QY 520 KTSQFVAPDELFGYTRKYSMAVMTG 545
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 544 ---YGRDAMAVGTPTKAVGVWVG 565
   | | | | | : : : | : : | : : | : : | : : | : : | : : | : : |
```

## RESULT 5

US-09-914-543-6  
; Sequence 6, Application US/09914543  
; GENERAL INFORMATION:  
; APPLICANT: DIVERSA CORPORATION  
; APPLICANT: LAM, David  
; APPLICANT: MATHUR, Eric  
; TITLE OF INVENTION: ENDOGLUCANASES  
; FILE REFERENCE: DIVER1150-5  
; CURRENT APPLICATION NUMBER: US/09/914,543  
; CURRENT FILING DATE: 1998-11-22  
; PRIOR APPLICATION NUMBER: PCT/US97/08793  
; PRIOR FILING DATE: 1997-05-22  
; PRIOR APPLICATION NUMBER: US 08/651,572  
; PRIOR FILING DATE: 1996-05-22  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 875  
; TYPE: PRN  
; ORGANISM: Bankia gouldi  
US-09-914-543-6

Query Match 4.6%; Score 162; DB 5; Length 875;  
Best Local Similarity 21.1%; Pred. No. 0.018;





```
Db 80 ROLDNT-----QTADQPKVMSDSATVKTSSN-----MQSPQANANOSTTKTS 127
Qy 119 NKVY--MSGNVYGHQTAQNY--YCKDLNNL-----SLPOLAL--LAGPQAPNOD 164
Db 128 NVTINDKSTSTYSNTDMSNLQAKDVSTTPKTTTIRKPRILRAAVNTVAAPQOSTNVND 187
Qy 165 P--YSHPEAODRRNLVLESEKNQGY--ISAEQYKAVNFTPTDGLQ-----207
Db 188 KVHFSNIDIAIDKGVHNVQTKTEFWATSSDYLKLANVYTIIDDSVKEGDTFFKYGQYER 247
Qy 208 --SLKSN--YPAVNDNLKEVINQVEEETGNLITGMDVYTNV-----249
Db 248 PGSVRLPQOTONLNAQGNIIAKGIYSTNTTYYT--FTNVYDQYTNVSGSEEQVAFKR 306
Qy 250 -----DOEAKHMDIYVTEDEVAAIPDELOASTIVDVNSNGVIAOLGARHOSNVFSG 304
Db 307 KNATTDKAYAMEVTLGN-DTY-----SEET-----IDVYGGKKAKOPLISSINYINNELDS 356
Qy 305 INQAVETNRDGSMTKP-----ITDY--APALEYGVYSTATTIVHDEPVNPGTNPVY 356
Db 357 RNMAYVAVQPKNTYTKQFVFNLTGCKFNPAKNKFIYEVI-----DQNGFVDSFTP-- 408
Qy 357 NWDGYSFQNTLQY--ALQOSRNPVAVETLNKVGILNRAKFTL-----NGLGID 402
Db 409 --DTSKLDVYDOFDVYISNDNKATVDMKQTSNKOYIIQVAYVDPNSSTDNG-KID 465
Qy 403 YP-----SIHYSAISNTTESD-KRYGASSEKMAAAYAFAGCY-----443
Db 466 YTLTDTKKYSWSNYSVNGSSSTANGDQKY-----NLGDTVWEDTKD 510
Qy 444 -----YKPMYIHKVYFSDGSEKEFSNVGRAMKETTAYMMTDMKTYLTYGTGNA 494
Db 511 GKQANKEGIGYV--VILKDSNGKELDRFTT--DENGKYOFTGLSNG--TYSVERST 562
Qy 495 YLALPQAGKIGTSNYDELENIHKTQFVAPD--ELFAGY--TRKYSMAVWTG-- 546
Db 563 PAGYPTTANVY--TDDAVSDGLTGTGVIKADANNLTDSGEFKYKPSYGLGYVWDS 618
Qy 547 -----SNRLPLVNGLTVAAYKVRSMYTYLSEGSPEDMNIPEG 586
Db 619 NKDGKRDSTENGIGVYKVTLONEKEGYIGTETDENGKTR--FDNLDSGKKVYFEKPAQ 676
Qy 587 LYRNG-----EFYFKNG--ARSTWNSPAPQPPSTESSSSS 621
Db 677 LTQGTNTTEDDKDADGGEVDYITLHDHDFLLDNGYEEFTSDSDSDSDSDSDSDSDS 736
Qy 622 DS-STOSSSTTPSTNNSTTNPNNNTQOSTTPDOONON 660
Db 737 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 776

RESULT 10
US-10-179-131-5286
; Sequence 3286, Application US/10179131
; GENERAL INFORMATION:
; APPLICANT: HARE, ROBERTA S.
; APPLICANT: SHAM, ROBERT J.
; APPLICANT: SHIMER JR., GEORGE H.
; APPLICANT: KESSLER, MARCO
; APPLICANT: NOLLING, JORR
; APPLICANT: ZENG, QIANDONG
; APPLICANT: GREENE, JONATHAN R.
; TITLE OF INVENTION: CANDIDA ALBICANS NUCLEIC ACIDS AND POLYPEPTIDES,
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: 2976-4031
; CURRENT APPLICATION NUMBER: US/10179,131
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 10194
; SEQ ID NO 5286
; LENGTH: 973
; TYPE: PRT
; ORGANISM: Candida albicans
```

```
US-10-179-131-5286
Query Match 4.4%; Score 152; DB 6; Length 973;
Best Local Similarity 19.0%; Pred. No. 0.099;
Matches 155; Conservative 91; Mismatches 312; Indels 258; Gaps 31;

Qy 78 KLYFSTSTSDQITSRKAQEMALAIQLEQATKQIELIYYINKYMSGNGYMQTAQN- 136
Db 109 EITLIGTSDTQSLNSNNYTMWVNSNDGLVLTQALLYSLSKSGQTFNGGLVVRPKD 168
Qy 137 -----YCKDLNLSLQ-----LALLAGMQANQVDP- 165
Db 169 IKIQEKNLFKSYSSDRPIIAYYGRSADRSSLNWIYFDGDTLTFSGTYVYVSENAPS 228
Qy 166 --YSHPEAODRRNLVLESEKNQGYISAEQYKAVN--TPRIDGLQSLKSNAYVAMDNY 222
Db 229 IDYSFSTIASDYGFAGAEKFKIYVGQHQLSTSNENETVNGITGKIDESTIILSDVF 288
Qy 223 L-----KEVINQVEEETGYMLT-----TGMDVYTNVDOEAKHMDIYVTEDEVYA 268
Db 289 LDGQPISEKENSIDYDQDLRYATFDKNNFTITGTFPTSTTDNFTIYVDIYGNVELP 348
Qy 269 YPDELOVASTI-----VDVSNKVIAQ-----LGARHOSNVSGINQAVETNNDWG 316
Db 349 YSFVYVNSIEFTIDSLKDVNATRGQYFQYQLKSYFTDAEDTKYVYDFGSG----TNSDW- 403
Qy 317 STMKPIIDYANALEYGVYSTATTIVHDEPVNPGTNPV--YVWDR-----360
Db 404 -----LOY--HDSMNTLSGITPKNFSLKVEINAESDSEKSFQIKGYDK 449
Qy 361 -----GYGNITLQYALQOSRNPVAVETLNKVGILNRAKFTLGLGIDYPSI 406
Db 450 IVSSSSSSSTSSSSSTASASSSFSSSSSP1AHKKS--NKNALALIGLGVIPVF 505
Qy 407 -----HYSNAISNTTESDKY-----GASEKMAA 433
Db 506 LILVAILLLOCCLRKRNKNENNTNDNEYENEFQPKRPMPIINGIAGAGAGAGA 565
Qy 434 YAAFANGGTYYKPMYIHKVYFSDGSEKEFSNVGRAM--KETTAYMMTDMKTYLTY 488
Db 566 LALAGAGAIY-----ATSSNESIKDESTMVFKLEHNTSKSSSITQVETSSSEFY 618
Qy 489 GTGNAYL--AMLPQA-----GKTGTSN-----YTDEIENHKTQFVA 526
Db 619 DTHENTPIVMSWRANTESDSKLARISGSLATVTEMLFVRLIDQYVNRSEISSEFLS 678
Qy 527 PDELFAGTTRYKSYMAVWTGYSNRLPLVNGGLIYA-----AKYRRSMAT--YISEG 575
Db 679 SNSLNALLRRESLSNNSNNSSTNFQRLDSNGNIYGELNPNNNNPKKNSRSSSEKMYTOL 738
Qy 576 SNP-----EDWN--IPEGLYRNGEYVFKNGARST-----WNS 605
Db 739 SSPLPPSLPQSGANLHIVPEKSRDLSHTGKDETTGTISNLLQFDNDSKYSVDEFFNS 798
Qy 606 PA-----POQPPSTESSSSS-----SDSST--SQSSST-----T 632
Db 799 SSRVKQQQQHQHPLSRNSFSNHVFDNRLPSPNVALGNSSTVWSPSSDFTLLDEST 858
Qy 633 PSTNN--STTNPNNNTQOSTTTPDOONONPQPAQ 665
Db 859 PINNHITNTTAPTATTNTNSNPNNNNNNDGQQTQ 894

RESULT 11
PCT-US02-19220-4
; Sequence 4, Application PC/TUS0219220
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.
; FILE REFERENCE: P072630US01/BAS
; CURRENT APPLICATION NUMBER: PCT/US02/19220
; CURRENT FILING DATE: 2002-06-17
```

; PRIOR APPLICATION NUMBER: US 60/298,098  
; PRIOR FILING DATE: 2001-06-15  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2283  
; TYPE: PRF  
; ORGANISM: Staphylococcus epidermidis  
; PCT-US02-19220-4

Query Match 4.3%; Score 148.5; DB 1; Length 2283;  
Best Local Similarity 20.6%; Pred. No. 0.57;  
Matches 143; Conservative 79; Mismatches 253; Indels 219; Gaps 31;

QY 63 SNSLOGS-----TLTQOLIKLTFSTSDQTSRKQAEAMIALIOLKATKQELITY 117  
DB 396 SNVAGGAGAFVTDSDYVATTTTSSSTADN-----AAKLNVQPTNPFQDPD 444  
QY 118 INKVMNNGN-----YGMFAON---YGRD-LNLSLPQALLAGMPQAPN-QYDP 165  
DB 445 IN-----YNDTKVMYKYGQTVTRNISWIAKSGTTNLSKTFASGATNIOQVFGT 500  
QY 166 YSHPEAADRNLVISEMKNQGYISAEOYEKAVNPEIT-DGIQSLKASANYPAYMDNYLK 224  
DB 501 FEYTESAVTQVRVY-DVTGKDIIPKTYSGNVYDVITIDNOQSALTAK----- 548  
QY 225 EYINQVEETGYNLTLTGMDVTNNVQEAQKHLMDITNDEVYAYVDEDLQVASTIVDVS 284  
DB 549 -----GYN-----YTSVDS-----SYASTYNDNTKTYKMTN 574  
QY 285 NKGVIQLGARHSSNVSEGINQAVETNRDMGSTMPI-----TDVAPALEYG 332  
DB 575 AQSQVYTYTVDKAPVYTG-NQTEYV-----GKTAMPYLTTDNGTGVTNTVGLPSG 629  
QY 333 V-YD-----STATVHDEPNYPGTN-----TPYVNDREG 362  
DB 630 LAYDSATNSIIGTPKIGQSTVTVSTDOANNKSTTTFTINVDYTAFTVTPIGDQSEV 689  
QY 363 FGNIT-LOYALQOSRNVPAVELINKVGLNRAKTF-----LNG-----LGIDYSHYSNA 411  
DB 690 YSPISPIKATODNSGNNAVNTVY-TGLPSGLFDSTNMTISGTPNIGSTISIVSTDA 747  
QY 412 ISSNTFESOKKYGASSEKMAAAYAFANGTYKKPMYIHKVYSDGSEKEFSVGRRAK 471  
DB 748 -GQNKTTTTFKTEVTRNSMDS-----VSTGSTQOSQSVSTSKAD 787  
QY 472 ETTAYMMTDMKTVLTYGTGRNAVYLAMLPQAGKTGTSNTDEIEHNIKTSQFVADPELF 531  
DB 788 SOSASTSTSGSIYVST-----SASTSKSTSVLSDSVASKSLSTSE-- 829  
QY 532 AGYTRKYSMAVNTGYSNRLTPLVGNGLVAAKYRSM-----MYLSEG-SNPEDMNIP 584  
DB 830 -----SNSYSSSTSTSLVNSQSVSSMSDSASKSTLSLSISNSSSTESK 874  
QY 585 EGL-----YRNGEFVFNNGARSTWNSPAPQOPSTESSSSSDS-----STSSQSSST 622  
DB 875 ESILSTSTSLKSTSLSDLSKSTSGLSKSTSLSTISGSSSTASLSLSDSTSNASTS 934  
QY 633 PSTNNSTTNPN-----NNTQOSNTTPOQN 658  
DB 935 TSLSESASTSDSISISNLSANQSASTSKSDSGS 968  
RESULT 12  
US-09-540-209B-6592  
; Sequence 6592, Application US/09540209B  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR  
; FILE REFERENCE: 2709.1001-001  
; CURRENT APPLICATION NUMBER: US/09/540,209B

; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 10444  
; SEQ ID NO 6592  
; LENGTH: 188  
; TYPE: PRF  
; ORGANISM: B.fragilis  
; US-09-540-209B-6592

Query Match 4.2%; Score 148; DB 5; Length 188;  
Best Local Similarity 27.6%; Pred. No. 0.018;  
Matches 48; Conservative 34; Mismatches 76; Indels 16; Gaps 6;

QY 38 EDHREFDHDGIDTRILGFLRLIQ-SNSLOGSTVTLQOLIKLTFSTSTSDQTSRKQAO 96  
DB 4 EDCATFFHKGFLPEARMKALYDLKARRAGSITTMOLVSVFLS---RKNIRAKLE 60  
QY 97 EA---WLAIOLEQKATKQELITYINKVYMSNGVMQTAQNYGKDLNLSLPQALL 153  
DB 61 EALIVWL-IETERTLSKERNYEVYLNIVEMGPLVGVQEAATYFPKRRPSQLAEESIFL 119  
QY 154 AGMPQAPNOY-----DPYSHPEAADRNLVISEMKNQGYIS--AEQYEKAVN 199  
DB 120 ASIIIPKPHFRNSFNNDMLKESLEGYRLITERLVKGIISEVADSIRPEIN 173

## RESULT 13

US-09-673-605A-34  
; Sequence 34, Application US/09673605A  
; GENERAL INFORMATION:  
; APPLICANT: The President and Fellows of Harvard College  
; TITLE OF INVENTION: REGULATION OF BIOFILM FORMATION  
; FILE REFERENCE: 00246/505003  
; CURRENT APPLICATION NUMBER: US/09/673,605A  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/102,870  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 60/083,259  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 595  
; TYPE: PRF  
; ORGANISM: Escherichia coli  
; US-09-673-605A-34

Query Match 4.2%; Score 147.5; DB 5; Length 595;  
Best Local Similarity 22.7%; Pred. No. 0.099;  
Matches 149; Conservative 89; Mismatches 291; Indels 127; Gaps 30;

QY 61 LOSNSLOGSSTLTLQOLIKLTFSTSTSDQTSRKQAEAMIAI-QLBOKATKQELITYIN 119  
DB 5 INTNSL---SLITQNNINNKQASLSSIERLS-----SGLRINSARDADAQAQIANRFTS 56  
QY 120 KV-----YMSNGYMGQTAQNTYG--KDLNLSLPQALLAGMPAPNOYDPYSHPEAA 172  
DB 57 NIKGLTQAARNANDGI-SVAQTEGALSETIN-NIQRIRBELT--VOASTGTNSDSDLS 112  
QY 173 QDRNLVISEMKNQGYISAEOYEKAVNPTITDLSLKSASNYPAYMDVYLKEVINQVEE 232  
DB 113 QDEIKRLEIDR---VSGOTQFNGVNVYLAQDSMAKIQVCANDGOTITITDKKIDSDTIG 169  
QY 233 ETGYNLTLTG-----MDVYTNVQEAQKHLMDITNDEVYVAYPD--DEL 274  
DB 170 LNFENNVNGSGTIANKAATISDLTAAKMDATNT-----ITTNALATSKALDOL 219  
QY 275 QVASTIVDVSNGKVIQLGARHSS--NVSEGINQAVETNRDMGSTMKPIITDVAPALEYGV 333  
DB 220 KQGDFTYIRADAQOTATVYTYNNAAGNFSFS--NVSNTSKAKADVAASLPLPAGQATASGV 278  
QY 334 YSTATIVHDEPNYPGTNTPYVNMDRGYFGNTLT--QVALQOSRNVPAVELINKVGLN 390

```
Db 279 YKASGEV-----NFVDANGKITIGGEAVLTSDGNL-----TTDAGGA 319
QY 391 RAKTFLNL---GIDYPSIHYSN--AISSNTTESDKKYGASSEKMAAAYAFANGCTYK 445
Db 320 TAAT-LDGLFKKAGGSGSIGFNKTASVTMGTTYNFKTGAD-----AGAATANG----- 368
QY 446 PMYIKVYFSDGSEKE--FSNVGTFRAMKETAYMTDMMKVTLYGTGRNAYLAMPQAG 503
Db 369 -----VSFTDASKETYLNKVAT--AKOGTAIVANGTSATITTKSGVQTYQAFAGD 420
QY 504 KTGISNTYDEETIENHIKTSQFVAPDEL--FAGYTRKYSMAVWTGY-----SNRLTPL 553
Db 421 GTASAKYADNTDVSNAATYTDADDEMTTIGSYTTRKYSIDANNGKVTVDGSGSKYAPK 480
QY 554 VGNGLTVAAKYRSMYTL-----SEGSNPD--WNIPEGLRNEEFYFKNGARSTWNSP 606
Db 481 VG-----AEVYVSANGTLTDTANSEGTVTKDPKLALDEAIISSIDKFSLSGLAIQNRLDS 534
QY 607 APQOPSTESSSSSDSTSOSTTPTSTNNSTTNPNNNQOQSWTTPDOONQO 662
Db 535 AVTNLNTNTTNLSEAKSRIQADADYATEVSNMS---KAQIIQAGNSVLAKANQVPO 587

RESULT 14
US-09-442-489B-2
; Sequence 2, Application US/09442489B
; GENERAL INFORMATION:
; APPLICANT: Albertsen, Hans
; APPLICANT: Anand, Rakesh
; APPLICANT: Carlson, Mary
; APPLICANT: Groden, Joanna
; APPLICANT: Hedge, Philip John
; APPLICANT: Josiyn, Geoff
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Markham, Alexander Fred
; APPLICANT: Nakamura, Yusuke
; APPLICANT: Thliveris, Andrew
; APPLICANT: Vogelstein, Bert
; APPLICANT: White, Raymond L.
; TITLE OF INVENTION: APC Antibodies
; FILE REFERENCE: 001107.78817
; CURRENT APPLICATION NUMBER: US/09/442.489B
; PRIOR FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: US 08/452,654
; PRIOR FILING DATE: 1995-05-25
; PRIOR APPLICATION NUMBER: US 08/289,548
; PRIOR FILING DATE: 1994-08-12
; PRIOR APPLICATION NUMBER: US 07/741,940
; PRIOR FILING DATE: 1991-08-08
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-442-489B-2

Query Match 4.2%; Score 146; DB 5; Length 2843;
Best Local Similarity 18.8%; Pred. No. 1.1;
Matches 130; Conservative 98; Mismatches 213; Indels 250; Gaps 31;
```

```
Db 801 ---FDNRHDDNRSDNFTGNMVTLPY-----LNTTV-----LPSSSSS 837
QY 216 PAYMDNYLKEVINOVEETQVNLTTGMDVYTNVQEAQKHLMDLYNDVEVYAPDDE-- 273
Db 838 RGSLSRSRSEKDRSLERBG-----IGLGNV-----HPATBNP 870
QY 274 -----LOVASTYVDVNGKVIQAOLGARHOSNVYFGINOAVETNRDMGST--MKPTTD 324
Db 871 GTSSKRGQISTTAAQIA--KMEVSAIHYS-----QEDRSSGTTTELHCYTD 917
QY 325 YAPALEGVYVSTATIVDEPYNTPGTTPVYNMDRGYFGNITLOYA--LOOSRNPVAVET 383
Db 918 ERNALR-----SSAAHTSNTYNF-----TKSENSNR-----TCSMPYAKLEKRS--SND 963
QY 384 LNKYGLN-----RAKTFNLGLID-----YPS-----IHYNAISSNTTESDK 421
Db 964 LNSVSSNDGYKRGOMKSISSYSEDDSKCSQYQYADLAHKIHSNHHMDNDGELDT 1023
QY 422 -----KYGASSEKMAAAYAFANGTYKPMYIHKVYFSDGSEKEFSNVGTFRAMKETTA 475
Db 1024 PINTSLKY--SDEQLNSGRQSPQONERAPRKHIIEDEIKQSEQOQSN-----QSTT 1074
QY 476 YMTDMKKTVLYTGGRNAYLAMPQAGKTGTSNTDEIENHIKTSQFVAPDELFCYGT 535
Db 1075 YPV-----YTESTDDKHLKQPHFQCECVSPYR 1103
QY 536 RYKSAVWTGYSNRLPLVYNGCLTVAAKYRSM-----TYSEGSNPDNWIPE 585
Db 1104 SR-----GANGSETNVGSHNHCINQVNSQSLCOEDDYEDDKPKNVSEKYSSEBQHEE 1156
QY 586 GLYRGEFVFRNGARSTWNSP-----APQOPSTESSSSSDSTSOSTT--PTNN 637
Db 1157 ERPTNYSIKYNEKRHV--DQIDYSLKATADIPSSQKQSFSSKSSSQSSQKTEHMSSS 1215
QY 638 STTNPNNTQOQN---TTPDOONQO 664
Db 1216 ENTSPSSNAKRONLHPSAQSRSQPOKA 1246
```

```
RESULT 15
US-10-179-131-6430
; Sequence 6430, Application US/10179131
; GENERAL INFORMATION:
; APPLICANT: HARE, ROBERTA S.
; APPLICANT: SHAW, KAREN J.
; APPLICANT: SHIMER JR., GEORGE H.
; APPLICANT: KESSLER, MARCO
; APPLICANT: NOLLING, JORR
; APPLICANT: ZENG, QIANDONG
; APPLICANT: GREENE, JONATHAN R.
; TITLE OF INVENTION: CANDIDA ALBICANS NUCLEIC ACIDS AND POLYPEPTIDES,
; AND USES THEREFOR
; FILE REFERENCE: 2976-4031
; CURRENT APPLICATION NUMBER: US/10/179.131
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 10194
; SEQ ID NO 6430
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-179-131-6430
```

```
Query Match 4.2%; Score 145.5; DB 6; Length 1029;
Best Local Similarity 18.9%; Pred. No. 0.29;
Matches 141; Conservative 106; Mismatches 233; Indels 267; Gaps 32;

QY 77 IKLYFSTSDQTSRRAQEWLAIOLEKATKOEILTYINKVYMGNGVYGMQTAQN 136
Db 264 IETATSSVADISIAKCSKNGPKPLKRTMEQSMISNLL-----KQLAQOI 316
QY 137 YYGKDLNLSLPOLALLAGMPQAPNOYDPSHPAADORRLVLSEMKNG--YISAQ 193
```



```
Db 317 HOKK--REOKLKOQEELKROHOL-NHPDEYDEEALSNKLN---SOYKNTAPYTHNTSVAK 370
QY 194 YEKAVNTPITDGLQSLKASANYPAYMDNYLKEVINQVEETG-----YN- 237
Db 371 LQSLKTPNSSSSASLKDLMKDEAVVPSEQISHQONQDGNVSGDYESKGERHIEHD 430
QY 238 --LTTGMDVYT-----NYDQE-----AQKHLMD-----IYN 262
Db 431 EVMQCIADIVYSDPEQRYNDEDEDYDSDDDDDYYDQYERPSNDSLAQSHLYEGDDEETEE 490
QY 263 TDEYVAVPDDE-----LOVASTIVDSNGKVIYAQLGARHOSNVSPGINQA 308
Db 491 ADEEV--EDEDEDEDEDEDEGGFFLVKSN---SNAPIT--LG-OHSASTSTPVAPS 541
QY 309 VETNRBMGSTMKPITD-----YAPALEYGVDS-----TATIVH 342
Db 542 LSRHTD-----ITDDTASISTTNSKSKYKTIOLEPSTSTNGSDSESDDEANPYTSSLH 594
QY 343 -----DEPYNYP-----GTNTPVY-----NMDRGYFGNITLQYALQ 374
Db 595 NVNNDISRGDYDYDNTFYTCNPNNSVYASQSPDVVDYPENLDMG--SNFDYEF-IEN 651
QY 375 SRVPAVEYLNKVGLNRAKTFGLNGLIDYPSIHYSAISSNTTESDKKYGASSEKMAAY 434
Db 652 NDSIPYVDI-----TFENNSTINNMPISYSP-----SPLSYAISG 688
QY 435 AAFANGTYYKPMYIHKVYFSDGSEKEFSNVGTRAKMETAYMMTDMKTVLTYGGRNA 494
Db 689 GGNKSGVTYVNSPNIYVNVNSNPQOQOQOQAKPKPTKASPQLSD----- 734
QY 495 YLAMLPOAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSMAVWTGYSNRLPLV 554
Db 735 -----SEDDSNSDSDDDGDISG-----LSIG-TRRSSQA----- 761
QY 555 GNGLTAAKAYRSMYTLSEGSNPEDMNIPGGLYRNGEYFKNGARSTWNSPAPQOPST 614
Db 762 ----LAESYFQSSLSSTQETAPQHHPDAKIEPAVEHVSINPR--YSSSTISKOPTS 814
QY 615 ESS-----SSSDSSTSOSSSTTPSTNNST----- 639
Db 815 SSSLISQSFEGAGLSSTKELSKSFLGSGTSASTSHDEKTTTIDSSSTGFGFOVPNRD 874
QY 640 -TTNPNNTQOQSNTPPDQONQNPQAPQ 665
Db 875 YTPSPDNNTLTFTLSNFSKKSPLPPQ 901
```

Search completed: August 11, 2002, 04:32:42  
Job time: 538 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 03:11:49 ; Search time 50.47 Seconds

(without alignments)  
1267.991 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KLYDNKNQLIADLGSERRVN.....TQSSNTTPDOQNPNQPAQP 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	719	2	G95042
2	3475	99.7	719	2	S28036
3	3471	99.6	719	2	S28037
4	3466	99.5	719	2	A42893
5	3466	99.5	719	2	A97913
6	3245	93.2	719	2	S28034
7	3245	93.1	719	2	S28032
8	3243	93.1	719	2	S28031
9	3241	93.0	719	2	S28033
10	3192	91.6	719	2	S28035
11	2820	80.9	608	1	S28036
12	2698.5	77.5	637	2	B42893
13	1657.5	47.6	664	2	G86692
14	1217	34.9	827	2	AD1311
15	1201	34.5	826	2	AD1683
16	1074.5	30.8	886	2	F83862
17	999	28.7	914	2	I40529
18	914	26.2	188	2	S31952
19	817.5	23.5	727	2	B89923
20	813.5	23.3	716	2	S43693
21	802.5	23.0	809	2	F97183
22	739.5	21.2	714	2	AE1353
23	739	21.2	714	2	AA0614
24	724	20.8	714	2	AF1723
25	712	20.4	726	2	F70355
26	698.5	20.0	719	2	A83800
27	670.5	19.2	643	2	E81396
28	656	18.8	966	2	E84053
29	655	18.8	764	2	B97371

30	655	18.8	764	2	AB2589	penicillin-binding
31	649.5	18.6	743	2	D86888	penicillin-binding
32	642.5	18.4	822	2	F83016	penicillin-binding
33	640	18.4	798	2	H81040	penicillin-binding
34	637	18.3	731	2	D95235	penicillin-binding
35	636	18.3	731	2	E98099	penicillin-binding
36	627	18.0	659	2	E64594	penicillin-binding
37	625	17.9	660	2	F71917	penicillin-binding
38	618	17.7	851	2	AF0018	penicillin-binding
39	617	17.7	835	2	B82051	penicillin-binding
40	615	17.7	755	2	A12174	penicillin-binding
41	614	17.6	873	2	B75514	penicillin-binding
42	613.5	17.6	687	2	D84126	penicillin-binding
43	608	17.5	718	2	A13420	penicillin-binding
44	598.5	17.2	624	2	A55220	penicillin-binding
45	598	17.2	885	2	S76357	penicillin-binding

## ALIGNMENTS

RESULT 1  
G95042 penicillin-binding protein 1A [imported] - Streptococcus pneumoniae (strain TIGR4)  
C.Species: Streptococcus pneumoniae  
C.Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001  
C.Accession: G95042

R.Tetelkin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H  
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001

A.Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris

A.Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A.Reference number: A95000; M01D:21357209; PMID:11463916

A.Accession: G95042

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-719 <KUD>

A.Cross-references: GB:AE005672; PIDN:AAK74536.1; PID:q14971838; GSPDB:GN00164; TIGR:

A.Experimental source: strain TIGR4

A.Genetics:

A.Gene: SP0369

C.Superfamily: penicillin-binding protein 1B

Query Match 100.0%; Score 3484; DB 2; Length 719;  
Best Local Similarity 100.0%; Pred. No. 4.6e-194;  
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KLYDNKNQLIADLGSERRVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLEFN	60
DB	54	KLYDNKNQLIADLGSERRVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLEFN	113
QY	61	LOSNSLOGSSTLTQOLIKLTFSTSTSDQTSIRKAOEAMLAQLBOKATKOEILTYINK	120
DB	114	LOSNSLOGSSTLTQOLIKLTFSTSTSDQTSIRKAOEAMLAQLBOKATKOEILTYINK	173
QY	121	VYMSNGNVMGMPTAQNYYKDLNNLSLPQALLAGMPQAPNDYDPYSHPDEAADRRLVL	180
DB	174	VYMSNGNVMGMPTAQNYYKDLNNLSLPQALLAGMPQAPNDYDPYSHPDEAADRRLVL	233
QY	181	SEMKNQGISAOYERKAVNTPITDGLQSLKSASNPAYADNLTKEVINYVEEETGNLLT	240
DB	234	SEMKNQGISAOYERKAVNTPITDGLQSLKSASNPAYADNLTKEVINYVEEETGNLLT	293
QY	241	TCMDVYTNVDQEAOKHLMIDIVMTDEVVAYPDELQVASTIVVNSNGKVAQAGARHSSN	300
DB	294	TCMDVYTNVDQEAOKHLMIDIVMTDEVVAYPDELQVASTIVVNSNGKVAQAGARHSSN	353
QY	301	VSEFGINQAVETNRDWSGTTKPTTDYAPALEYGVYDSTATIVHDEPNKPGINTPYVNMNR	360
DB	354	VSEFGINQAVETNRDWSGTTKPTTDYAPALEYGVYDSTATIVHDEPNKPGINTPYVNMNR	413

Qy	361	GYFGNITLQYALQOSRNPAYETLANKYGLNRKATFLNGLGIDPSTHYSAISSNTTESD	420
Db	414	GYFGNITLQYALQOSRNPAYETLNKYGILNRKATFLNGLGIDPSTHYSAISSNTTESD	473
Qy	421	KYGAASSEKMAAAYAANGSTYYKPMYIHKVVSDDSEKEFSVNGVRAKETTYAVMTD	480
Db	474	KYGAASSEKMAAAYAANGSTYYKPMYIHKVVSDDSEKEFSVNGVRAKETTYAVMTD	533
Qy	481	MKTYLVTYGTGRNAYLALPQAGKTGTSNYDEIEIHNKTSQFVAPDELFCAGYTRKYSM	540
Db	534	MKTYLVTYGTGRNAYLALPQAGKTGTSNYDEIEIHNKITSQFVAPDELFCAGYTRKYSM	593
Qy	541	AVMTGYSNRLTPLYVNGSLTVAKYVRSMTYILSEGSNPEDMNIDEGLYRNGEEVFVKNGAR	600
Db	594	AVMTGYSNRLTPLYVNGSLTVAKYVRSMTYILSEGSNPEDMNIDEGLYRNGEEVFVKNGAR	653
Qy	601	STWNSPAPQPPSTBESSSSSSDSTSOSSSTTPTNNSTTTPNPNNTQOSNTTTPDOONON	660
Db	654	STWNSPAPQPPSTBESSSSSSDSTSOSSSTTPTNNSTTTPNPNNTQOSNTTTPDOONON	713
Qy	661	POPAPQ 666	
Db	714	POPAPQ 719	

RESULT 2  
S28038  
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 45607) (fragment)  
C:Species: Streptococcus pneumoniae  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 26-May-2000  
C:Accession: S28038  
R:Martin, C.; Sibold, C.; Hakenbeck, R.  
EMBO J. 11, 3831-3836, 1992  
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant strains of *Streptococcus pneumoniae*  
A:Reference number: S28031, MUID:93010977  
A:Accession: S28038  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-719 <MAP>  
A:Cross-references: EMBL:X67873; NID:g47419; PIDN:CAA6073.1; PID:g47420  
C:Superfamily: penicillin-binding protein 1B

Query Match	99.7%	Score 34/75	DB 2	Length 719
Best Local Similarity	99.7%	Pred. No. 1.5e-193		
Matches 664	Conservative	2	Mismatches 0	Indels 0
QY	1 KTYDDKNQDLIADLGSERRRNQAQNDIPDVLKATYSIEDHREFDRGIDITIRILGAFLRN	60		
Db	54 KTYDDKNQDLIADLGSERRRNQAQNDIPDVLKATYSIEDHREFDRGIDITIRILGAFLRN	113		
QY	61 LOSNLSLOGGSTLTQOOLIKLTFPSTSTSDPTISRRKQEAAMLALQLOKATKOEILYYIYK	120		
Db	114 LOSNLSLOGGSTLTQOOLIKLTFPSTSTSDPTISRRKQEAAMLALQLOKATKOEILYYIYK	173		
QY	121 YVMSNGNNGMGTAAQNYGKDLNNLSLDPQALLACGPPQAPQNDYDEYSHPBEAQQDRNNLVL	180		
Db	174 YVMSNGNNGMGTAAQNYGKDLNNLSLDPQALLACGPPQAPQNDYDEYSHPBEAQQDRNNLVL	233		
QY	181 SEMKNGYITSAEQYKKAQVNTPTDGLQSLKSAQNPAYMDNYLKEYINQVEEETGYNLT	240		
Db	234 SEMKNGYITSAEQYKKAQVNTPTDGLQSLKSAQNPAYMDNYLKEYINQVEEETGYNLT	293		
QY	241 TGMDDVYTVNDDQAQKHLMDIYNTDERVAYRPDELOVASTYDVVSGKYTAQLGARHOSN	300		
Db	294 TGMDDVYTVNDDQAQKHLMDIYNTDERVAYRPDELOVASTYDVVSGKYTAQLGARHOSN	353		
QY	301 VSEFGINQAVERTRBDGSPMKPTTDAPALPEYGVYOSTATVYDEPNYPGNTPTVYNDMR	360		
Db	354 VSEFGINQAVERTRBDGSPMKPTTDAPALPEYGVYOSTATVYDEPNYPGNTPTVYNDMR	413		
QY	361 GFYGNITLQYALAQQRNPAVETLLKVGILNRAKTEPLANGIGIDYPSIIHSNATISMTTESD	420		

Db	414	GFGNITLQIYALQOSRNPVAVETLANKGILNRAKTFLENGLGIDRYSIHYSNAISNTTESD	473
Qy	421	KKYGASSEKMAAAYAAEANGSTYYKPMYIHKVPSDSEKEFEFNVGRAMEKETTAYMTD	480
Db	474	KKYGASSSEKMAAAYAAEANGSTYYKPMYIHKVVPDSDEKEFEFNVGRAMEKETTAYMTD	533
Qy	481	MMKTLVLTGTGGRNATLAWLPAGKGTGTSNTYDEIEIENHITKSQFVAPDELFACTYRYSM	540
Db	534	MMKTLVLTGTGGRNATLAWLPAGKGTGTSNTYDEIEIENHITKSQFVAPDELFACTYRYSM	593
Qy	541	AVMTGYSNRLPLVYNGELITVAAKYRSMYTLSEGSNPEDMNIPEGLYRNGEYEFKNGAR	600
Db	594	AVMTGYSNRLPLVYNGELITVAAKYRSMYTLSEGSNPEDMNIPEGLYRNGEYEFKNGAR	653
Qy	601	STWNPSAPQOQPPSTFESSSSSSDSSSTSSSTTPTNNSTTTPNPNNTQOSNTTPDOONON	660
Db	654	STWNPSAPQOQPPSTFESSSSSSDSSSTSSSTTPTNNSTTTPNPNNTQOSNTTPDOONON	713
Qy	661	POPAPQ 666	
Db	714	POPAPQ 719	

RESULT 3  
S28037  
penicillin-binding protein Ia - Streptococcus pneumoniae (strain 63915) (fragment)  
C.Species: Streptococcus pneumoniae  
C.Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 26-May-2000  
C.Accession: S28037  
R.Martin, C.; Sibold, C.; Hakenbeck, R.  
EMBO J. 11, 3831-3836, 1992  
A.Title: Relatedness of penicillin-binding protein Ia genes from different clones of  
A.Reference number: S28031; MUID:93010977  
A.Accession: S28037  
A.Status: translation not shown  
A.Molecule type: DNA  
A.Residues: 1-719 <MAR>  
A.Cross-references: EMBL:X67872; NID:g47417; PIDN:CAA48072.1; PID:g47418  
C.Superfamily: penicillin-binding protein 1B

Query Match	99.6%	Score 3471	DB 2	Length 719
Best Local Similarity	99.4%	Pred. No. 2,66-193		
Matches 662	Conservative 4	Missed matches 0	Indels 0	Gaps 0
QY	1	KIYDKNOLIALGSERRYNAOANIPMDLYKAISIEDHRRFDHGGIDTIRILGAFLRN	60	
Db	54	KIYDKNOLIALDGSERRYNAOANIPMDLYKAISIEDHRRFDHGGIDTIRILGAFLRN	113	
QY	61	LOSNSLOGSGSTLTQOLIKLTYFSTSTSDQTSIRKAOEAMLAIOLEOKATKOEILTYYNK	120	
Db	114	LOSNSLOGSGSTLTQOLIKLTYFSTSTSDQTSIRKAOEAMLAIOLEOKATKOEILTYYNK	173	
QY	121	YMSNGNNGMCPAANYYGKDLNNLSLPQALLAAMPQAPNPDYSHPEAODRRNLVL	180	
Db	174	YMSNGNNGMCPAANYYGKDLNNLSLPQALLAAMPQAPNPDYSHPEAODRRNLVL	233	
QY	181	SEMKNOGYISAEQYKAAVNTPTDLOSLKSASNPAYMDNYLKEVINQVEBETGYNLT	240	
Db	234	SEMKNOGYISAEQYKAAVNTPTDLOSLKSASNPAYMDNYLKEVINQVEBETGYNLT	293	
QY	241	TGMDVYTVNDGEAKNHLMDIYNTDEYVAYRPDELOVASTIYDVSNGKYIAOLGARHOSSN	300	
Db	294	TGMDVYTVNDGEAKNHLMDIYNTDEYVAYRPDELOVASTIYDVSNGKYIAOLGARHOSSN	353	
QY	301	VSEFGINQAVETNRDGSSTMKPTTADAPALEGVYUUSTATIVHDEEYANPGTNTPYNMNR	360	
Db	354	VSEFGINQAVETNRDGSSTMKPTTADAPALEGVYUUSTATIVHDKRYNPGTNTPYNMNR	413	
QY	361	GYEGNITLOVALQOOSRNPVAVETLNKVGILNRAKTFPLNGLGIDYPSIHSNMAISSWTESD	420	
Db	414	GYEGNITLOVALQOOSRNPVAVETLNKVGILNRAKTFPLNGLGIDYPSIHSNMAISSWTESD	473	

421 KKYGASSEKMAAAYAFANGGTYRKPMYIHKVVFSDGSEKESNVGTRAMKETTAAMMD 480  
|||||  
474 KKYGASSEKMAAAYAFANGGTYRKPMYIHKVVFSDGSEKESNVGTRAMKETTAAMMD 533  
Db  
481 MKKTVLVTGTGRNAYLAWLPQAGKTGTSNYTDEEIEENHIKTSQFVAPDELFGAGYTRKYSM 540  
|||||  
534 MKKTVLSYGTGRNAYLAWLPQAGKTGTSNYTDEEIEENHIKTSQFVAPDELFGAGYTRKYSM 593  
Db  
541 AVMTGYSNRLLTPVLVNGGLVAAKYVRSMMTYLSSEGSNPEDMNIPGLYRNGEYFVKNGAR 600  
|||||  
594 AVMTGYSNRLLTPVLVNGGLVAAKYVRSMMTYLSSEGSNPEDMNIPGLYRNGEYFVKNGAR 653  
Db  
601 STMWSPAPQPPSTESSSSSDSTSSQSSSTTPTNNSTTTTPNNNTQOOSNTTTPQOONON 660  
|||||  
654 STMWSPAPQPPSTESSSSSDSTSSQSSSTTPTNNSTTTTPNNNTQOOSNTTTPQOONON 713  
Db  
661 POPAQP 666  
|||||  
714 POPAQP 719  
Db

RESULT 4  
A42893  
penicillin-binding protein 1A - Streptococcus pneumoniae  
C:Species: Streptococcus pneumoniae  
C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 21-Jan-2000  
C:Accession: A42893  
R:Martin, C.; Briese, T.; Hakenbeck, R.  
J. Bacteriol. 174, 4517-4523, 1992  
A:Title: Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus pneumoniae.  
A:Reference number: A42893; MUID:92325042  
A:Accession: A42893  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-719 <MAR>  
A:Cross-references: GB:M0527; NID:g153766; PIDN:AAA26956.1; PID:g153768  
C:Superfamily: penicillin-binding protein 1B

Query Match 99.5%; Score 3466; DB 2; Length 719;  
Best Local Similarity 99.4%; Pred. No. 5.1e-193;  
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

421 KKYGASSEKMAAAYAFANGGTYRKPMYIHKVVFSDGSEKESNVGTRAMKETTAAMMD 480

Db  
474 KKYGASSEKMAAAYAFANGGTYRKPMYIHKVVFSDGSEKESNVGTRAMKETTAAMMD 533  
Db  
481 MKKTVLVTGTGRNAYLAWLPQAGKTGTSNYTDEEIEENHIKTSQFVAPDELFGAGYTRKYSM 540  
|||||  
534 MKKTVLSYGTGRNAYLAWLPQAGKTGTSNYTDEEIEENHIKTSQFVAPDELFGAGYTRKYSM 593  
Db  
541 AVMTGYSNRLLTPVLVNGGLVAAKYVRSMMTYLSSEGSNPEDMNIPGLYRNGEYFVKNGAR 600  
|||||  
594 AVMTGYSNRLLTPVLVNGGLVAAKYVRSMMTYLSSEGSNPEDMNIPGLYRNGEYFVKNGAR 653  
Db  
601 STMWSPAPQPPSTESSSSSDSTSSQSSSTTPTNNSTTTTPNNNTQOOSNTTTPQOONON 660  
|||||  
654 STMWSPAPQPPSTESSSSSDSTSSQSSSTTPTNNSTTTTPNNNTQOOSNTTTPQOONON 713  
Db  
661 POPAQP 666  
|||||  
714 POPAQP 719  
Db

RESULT 5  
A97913  
peptidoglycan glycosyltransferase (EC 2.4.1.129) [imported] - Streptococcus pneumonia  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 02-Nov-2001  
C:Accession: A97913  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.  
e, R.; LeBlanc, D.J.; Lee, L.N.; Letkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: A97913  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-719 <KUR>  
A:Cross-references: GB:AEO07317; PIDN:AAK9133.1; PID:g15457886; GSPDB:GN00174  
C:Genetics:  
A:Gene: pbpa  
C:Superfamily: penicillin-binding protein 1B  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 99.5%; Score 3466; DB 2; Length 719;  
Best Local Similarity 99.4%; Pred. No. 5.1e-193;  
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

421 KKYGASSEKMAAAYAFANGGTYRKPMYIHKVVFSDGSEKESNVGTRAMKETTAAMMD 480

```
|||||
Db 414 GYFNGITLQYALQOOSRNPVAVETLNKGLNRAKFFLNGLDIDPSIHNSNISSNTTSD 473
QY 421 KRYGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKFSNVGTTRAMKETTAYMTD 480
Db 421 KRYGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKFSNVGTTRAMKETTAYMTD 480
Db 474 KRYGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKFSNVGTTRAMKETTAYMTD 533
QY 481 MKKTVLYTGTGRNAVLAFLPQAGKGTGTSNYTDEEIEHNKTSQFVAPDELFAGYTRKYSM 540
Db 534 MKKTVLYTGTGRNAVLAFLPQAGKGTGTSNYTDEEIEHNKTSQFVAPDELFAGYTRKYSM 593
QY 541 AVMTGYSNRLLPVLGNGLTVAAKYRSMYTLSEGSNPEDNNIPEGLRNGEFYFKNGAR 600
Db 594 AVMTGYSNRLLPVLGNGLTVAAKYRSMYTLSEGSNPEDNNIPEGLRNGEFYFKNGAR 653
QY 601 STWNSPAPQOPSTRESSSSSSDSSSTPSTPSTNNSTTNPNNNTQOQSNTPPDQOQN 660
Db 654 STWNSPAPQOPSTRESSSSSSDSSSTPSTPSTNNSTTNPNNNTQOQSNTPPDQOQN 713
QY 661 POPAOP 666
Db 714 POPAOP 719
```

```
RESULT 6
S28034
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 56742) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28034
R:Martin, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of per
A:Reference number: S28031; MUID:93010977
A:Accession: S28034
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67869; NID:947411; PIDN:CAA48069.1; PID:947412
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 93.28; Score 3246; DB 2; Length 719;
Best Local Similarity 92.08; Pred. No. 2.8e-180;
Matches 613; Conservative 31; Mismatches 22; Indels 0; Gaps 0;
QY 1 KIYNKNKQNLINDLSERRVNAQANDIPTDLKATYSIEDHFPDRGIDTIRILGAFLRN 60
Db 54 KIYNKNKQNLINDLSERRVNAQANDIPTDLKATYSIEDHFPDRGIDTIRILGAFLRN 113
QY 61 LOSNSLOGGSTLTQOLIKLTYFSTSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 120
Db 114 LOSNSLOGGSTLTQOLIKLTYFSTSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 173
QY 121 VYNSNGNYGMOTAAQNYGKDLNLSLPQALLAGMPAPQOYPSHPEAQAODRRNLVL 180
Db 174 VYNSNGNYGMOTAAQNYGKDLNLSLPQALLAGMPAPQOYPSHPEAQAODRRNLVL 233
QY 181 SEMNNOGYSIEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVQEGYNYLT 240
Db 234 SEMNNOGYSIEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVQEGYNYLT 293
QY 241 TGMVYTNVDEQAKHLMIDINTDEYVAYPDDELQVASTIYDVNSGKVIQALGARHOSSN 300
Db 294 TGMVYTNVDEQAKHLMIDINTDEYVAYPDDELQVASTIYDVNSGKVIQALGARHOSSN 353
QY 301 VSFGINQAVETNRDWSGSMKPIITDYAPALEGYVDSTATIVHDEPYNPQGTNPYVYMDR 360
Db 354 VSFGINQAVETNRDWSGSMKPIITDYAPALEGYVDSTATIVHDEPYNPQGTNPYVYMDR 413
QY 361 GYFNGITLQYALQOOSRNPVAVETLNKGLNRAKFFLNGLDIDPSIHNSNISSNTTSD 420
Db 414 GYFNGITLQYALQOOSRNPVAVETLNKGLNRAKFFLNGLDIDPSIHNSNISSNTTSD 473
```

```
QY 421 KRYGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKFSNVGTTRAMKETTAYMTD 480
Db 474 KRYGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKFSNVGTTRAMKETTAYMTD 533
QY 481 MKKTVLYTGTGRNAVLAFLPQAGKGTGTSNYTDEEIEHNKTSQFVAPDELFAGYTRKYSM 540
Db 534 MKKTVLYTGTGRNAVLAFLPQAGKGTGTSNYTDEEIEHNKTSQFVAPDELFAGYTRKYSM 593
QY 541 AVMTGYSNRLLPVLGNGLTVAAKYRSMYTLSEGSNPEDNNIPEGLRNGEFYFKNGAR 600
Db 594 AVMTGYSNRLLPVLGNGLTVAAKYRSMYTLSEGSNPEDNNIPEGLRNGEFYFKNGAR 653
QY 601 STWNSPAPQOPSTRESSSSSSDSSSTPSTPSTNNSTTNPNNNTQOQSNTPPDQOQN 660
Db 654 PIMTEPSTQOQSTRESSSSSSDSSSTPSTPSTNNSTTNPNNNTQOQSNTPPDQOQN 713
QY 661 POPAOP 666
Db 714 POPAOP 719
```

```
RESULT 7
S28032
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 681) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28032
R:Martin, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MUID:93010977
A:Accession: S28032
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67866; NID:947407; PIDN:CAA48066.1; PID:947408
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 93.18; Score 3245; DB 2; Length 719;
Best Local Similarity 92.08; Pred. No. 3.2e-180;
Matches 613; Conservative 35; Mismatches 18; Indels 0; Gaps 0;
QY 1 KIYNKNKQNLINDLSERRVNAQANDIPTDLKATYSIEDHFPDRGIDTIRILGAFLRN 60
Db 54 KIYNKNKQNLINDLSERRVNAQANDIPTDLKATYSIEDHFPDRGIDTIRILGAFLRN 113
QY 61 LOSNSLOGGSTLTQOLIKLTYFSTSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 120
Db 114 LOSNSLOGGSTLTQOLIKLTYFSTSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 173
QY 121 VYNSNGNYGMOTAAQNYGKDLNLSLPQALLAGMPAPQOYPSHPEAQAODRRNLVL 180
Db 174 VYNSNGNYGMOTAAQNYGKDLNLSLPQALLAGMPAPQOYPSHPEAQAODRRNLVL 233
QY 181 SEMNNOGYSIEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVQEGYNYLT 240
Db 234 SEMNNOGYSIEQYKAVNPITDGLQSLKSASNPAYMDNYLKEVINQVQEGYNYLT 293
QY 241 TGMVYTNVDEQAKHLMIDINTDEYVAYPDDELQVASTIYDVNSGKVIQALGARHOSSN 300
Db 294 TGMVYTNVDEQAKHLMIDINTDEYVAYPDDELQVASTIYDVNSGKVIQALGARHOSSN 353
QY 301 VSFGINQAVETNRDWSGSMKPIITDYAPALEGYVDSTATIVHDEPYNPQGTNPYVYMDR 360
Db 354 VSFGINQAVETNRDWSGSMKPIITDYAPALEGYVDSTATIVHDEPYNPQGTNPYVYMDR 413
QY 361 GYFNGITLQYALQOOSRNPVAVETLNKGLNRAKFFLNGLDIDPSIHNSNISSNTTSD 420
Db 414 GYFNGITLQYALQOOSRNPVAVETLNKGLNRAKFFLNGLDIDPSIHNSNISSNTTSD 473
QY 421 KRYGASSEKMAAAYAAAFANGCTYYKPMYIHKVPSDGSSEKFSNVGTTRAMKETTAYMTD 480
```

Db	474	KÖYÖSSSEKMAAAYAAAFANGSIYHKPNTINKVPSDOSKKEFEVDGRRAKETTAAHME	533
Qy	481	MAKTYLYTGTORRAAYLAWLPQAGKTGTSNTYDDETEHNIKTSCQVAPDELFAGYTKRYS	540
Db	534	MAKTYLVAAGTGRGAYLLEPMLAORGKTGTSNTYDDETEHNIKTSCYVADEMEFYTKRYS	593
Qy	541	AWMTGYSNRLLPIANGNLTYAAKYRSMATYLSGCSNPEDMNTPEGLYRNGEEVFNKGAR	600
Db	594	AWMTGYSNRLLPIYGDGGLYAAKYRSMITYLSDETTPEDMTMPDGLFRNGEEVFNKGAR	653
Qy	601	STWNSPAPOQPPSTRESSSSSSSDSTSQSSSTPTNNSTTNNNNNQSSNTTPODONON	660
Db	654	STWNSPAPOQPPSTRESSSSSSSDSTSQSSSTPTNNSTTDDNNNTQSSNTTTPDONON	713
Qy	661	POPAOP 666	
Db	714	POPAOP 719	

RESULT 8  
S28031  
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 456) (fragment)  
C:Species: Streptococcus pneumoniae  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 26-May-2000  
C:Accession: S28031  
R:Martin, C.; Sibold, C.; Hakenbeck, R.  
EMBO J. 11, 3831-3836, 1992  
A>Title: Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant strains of *Streptococcus pneumoniae*  
A:Reference number: S28031; MUID:93010977  
A:Accession: S28031  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-719 <MAR>  
A:Cross-references: EMBL:X67868; NID:q47405; PIDN:CAA48068.1; PID:q47406  
C:Superfamily: penicillin-binding protein 1a

Query Match	93.1%	Score 3243:	DB 2	Length 719:
Best Local Similarity	91.9%:	Pred. No. 4.2e-180:		
Matches 612:	Conservative 32:	Mismatches 22:	Indels 0:	Gaps 0:
QY	1 KIYDNKNOLIADLGSERRYNAQANDIPITDLYAKAIVSIEDHREFDHKGIDITIRILGAFLRN	60		
DB	54 KIYDNKNOLIADLGSERRYNAQANDIPITDLYAKAIVSIEDHREFDHKGIDISIRILGAFLRN	113		
QY	61 LOSNLSGGSGSTLTQOLIKLITYFSTSTSDOTISRKQEAHLALQLODKAKIOLITYYINIK	120		
DB	114 LOSNLSGGSGSTLTQOLIKLITYFSTSTSDOTISRKQEAHLALQLODKAKIOLITYYINIK	173		
QY	121 VYMSNGNGYMGQTAANYYAKDKANLNLSPLOALLAGMPQAPNOPYDPSHPEAADRRLNL	180		
DB	174 VYMSNGNGYMGQTAANYYAKDKANLNLSPLOALLAGMPQAPNOPYDPSHPEAADRRLNL	233		
QY	181 SEMKNOGYISAOYEKAVNPTITDGLQSLKSASNPAYNDNLKEYINOVEETGYNLT	240		
DB	234 SEMKNOGYISAOYEKAVNPTITDGLQSLKSASNPAYNDNLKEYINOVEETGYNLT	293		
QY	241 TGMDEVYTNVDCQAKHLMIDYNTDEYVAPDDELQVASTIVYVNSNGKVAIQAQCARHQSNN	300		
DB	294 TGMDEVYTNVDCQAKHLMIDYNTDEYVAPDDELQVASTIVYVNSNGKVAIQAQCARHQSNN	353		
QY	301 VSEGTINQAVETNRDMSGSTMKPTITDVAPEALEYGVYDSTATIVHDEPNYNGTNTPYVNMDR	360		
DB	354 VSEGTINQAVETNRDMSGSAKPTITDVAPEALEYGVYDSTATIVHDEPNYNGTSTPYVNMDR	413		
QY	361 GYEGNTITLOALQOOSNNVPAVEITLKNVGNRKAFTFNGTIDYPSYIHANALISSNTTSEN	420		
DB	414 AVEGNTITLOALQOOSRNTAVAVETLKNVGLDRAKFTFNGTIDYPSMHYANALISSNTTSEN	473		
QY	421 KRYGASSEKMAAAYAFANGGYTYKPMYTHKVVSDGSEKESNVGTRAMKETAYAMMD	480		
DB	474 KRYGASSEKMAAAYAFANGGYTYKPMYTHKVVSDGSEKESNVGTRAMKETAYAMMD	533		

QY	481	MMPLVLTGSTRNNYLA	LPQAGTGTGNSNYNDE	IEHHIT	TSQVAPDEL	PACTYTRYSM	540	
Db	534	MKKVLTLYGTGSGAL	PLPLPQAGTGTGNSNT	DEEVHNI	KNTGVAPDEMF	VGITYTRYSM	593	
QY	541	AVMTGYSNRLTPIV	ONGILTVAAKYRSM	MTYLSG	SNPEDMNI	PEGLRYNGEFVFKGAR	600	
Db	594	AVMTGYSNRLTPIY	GDGTLVAAKYRS	MTYLS	EDTPEDMTMDGL	FRNGEVEVFKGAR	653	
QY	601	STWNSPAPQPPST	BSRSSSSD	STSQSS	STTPT	NNSTTTPNPNNT	QOQSNTPDOONON	660
Db	654	PIWTPSTQOQST	AESSSSSSD	STSQSS	STTPT	NNSTTTPNPNNT	QOQSNTPDOONON	713
QY	661	POPAPQ	666					
Db	714	POPAPQ	719					

RESULT 9  
 S28033  
 penicillin-binding protein 1a - Streptococcus pneumoniae (strain 670) (fragment)  
 C.Species: Streptococcus pneumoniae  
 C.Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 26-May-2000  
 C.Accession: S28033  
 R.Martin, C.; Sibold, C.; Hakenbeck, R.  
 EMBO J. 11, 3831-3836, 1992  
 A.Title: Relatedness of penicillin-binding protein 1a genes from different clones of  
 A.Reference number: S28031; MUID:93010977  
 A.Accession: S28033  
 A.Status: translation not shown  
 A.Molecule type: DNA  
 A.Residues: 1-719 <MAR>  
 A.Cross-References: EMBL:X67867; NID:q47409; PIDN:CAA48067.1; PID:q47410  
 C.Superfamily: penicillin-binding protein 1B

Query Match	93.0%	Score 3241	DB 2241	length 719
Best Local Similarity	91.9%	Pred. No. 5.5e-180		
Matches 612	Conservative	36	Mismatches 18	Indels 0
				Gaps 0
QY	1	KTYDNKNOLIALDGSRRRYNAOANDIPDTLVAAIYSIEDHREFDHRGIDTIRILGAFLELN	60	
DB	54	KTYDNKNOLIALDGSRRRYNAOANDIPDTLVAAIYSIEDHREFDHRGIDISILILGAFLELN	113	
QY	61	IQSNLSLGGSTLTLOOLIKLTFYESTSDQTISRKAQEAALIAQLBQAKATKOEITLYYINK	120	
DB	114	IQSNLSLGGSTLTLOOLIKLTFYESTSDQTISRKAQEAALIAQLBQAKATKOEITLYYINK	173	
QY	121	VYMSNGNGYMOGTAOANYGYKDLNNISLPOLALLAGHPQAPNOYDPSHPEAADRNNLYL	180	
DB	174	VYMSNGNGYMOGTAOANYGYKDLNNISLPOLALLAGHPQAPNOYDPSHPEAADRNNLYL	233	
QY	181	SEMKNNGYISARQYEKAVNPTITDGLQSLKASNPAYMDNTLKEYNVEEETGNILTL	240	
DB	234	SEMKNNGYISARQYEKAVNPTITDGLQSLKASNPAYMDNTLKEYNVEEETGNILTL	293	
QY	241	TEMDYVYNTDQEAOKHLMIDYNTDEVAAPDDELQVASTIVYVNSNGKYTAQGLGARHOSN	300	
DB	294	TEMDYVYNTDQEAOKHLMIDYNTSDQVYSTPDDDLQVASTIVYVNSNGKYTAQGLGARHOSN	353	
QY	301	VSEGINQAVETNRDMSGTMRPTTDYAPALEGYVDSTATIVHDEPNYNGTNPVYNNMDR	360	
DB	354	VSEGINQAVETNRDMSGAMKPTTDYAPALEGYVDSTATVMNDIPNPGTSPMYNNMDR	413	
QY	361	GTFGNITLQYALQOOSNNVAVETLNKVLGNRAKFTPLNGIGIDYPSLHYSNAISSNTSEED	420	
DB	414	AFFGNITLQYALQOOSNNVAVETLNKVLGNRAKFTPLNGIGIDYPSKHYANAISSNTSEEN	473	
QY	421	KYVGASSEKMAAAYAAFAFGGYTPMYTHKVVESQSGSEKESNNGTRBMKFTTYAAMMD	480	
DB	474	KYVGASSEKMAAAYAAFAFGGYTHKPMYTHKVVESQSGSEKESSDGTRBMKFTTYAAMTE	533	
QY	481	MKKTYLVYGTGNAYLAWLPQAGKGTGSNTVDEEIEHHKITSQFVAPADELFAGYTRKYSM	540	

```
Db 534 MKKTVLAGTGTGAGALPMLPAQAGKTGTSNYTDELEIKHKNKGVAPEMFGYTRKYSM 593
OY 541 AVMTCYSNRLLPVLVNGTLVAAKYVRSMWTYLSGSGNPEDNMIPEGGLRNGEFVKNNGAR 600
Db 554 AVMTCYSNRLLPVLVNGTLVAAKYVRSMWTYLSGSGNPEDNMIPEGGLRNGEFVKNNGAR 653
OY 601 STWNSPAPQOQPPSTRESSSSSDSSSTPSTNNSTTTNNPNNNQOOSNTTTPDOQONN 660
Db 654 STWNSPAPQOQPPSTRESSSSSDSSSTPSTNNSTTTNNPNNNQOOSNTTTPDOQONN 713
OY 661 POPAQP 666
Db 714 POPAQP 719
```

## RESULT 10

S28035  
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 2039) (fragment)

C:Species: Streptococcus pneumoniae

A:Variety: strain 2039

C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 26-May-2000

C:Accession: S28035

R:Martin, C.; Sibold, C.; Hakenbeck, R.

EMBO J. 11, 3831-3836, 1992

A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of pen

A:Reference number: S28031; MUID:93010977

A:Accession: S28035

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-719 <MAR>

A:Cross-references: EMBL:X67870; NID:q47413; PIDN:CAA48070.1; PID:q47414

A:Experimental source: strain 2039

C:Superfamily: penicillin-binding protein 1B

C:Keywords: antibiotic resistance; penicillin resistance

Query Match 91.6%, Score 3192; DB 2; Length 719;  
Best Local Similarity 90.4%, Pred. No. 3.8e-177;  
Matches 602; Conservative 41; Mismatches 23; Indels 0; Gaps 0;

```
OY 1 KIYDKKNQIADLGSERRVNAQANDIPDLVKAIVSIEDHREFDRGIDITRIILGAFRLN 60
Db 54 KIFDSKNELIADLGSERRVNAQANDIPDLVKAIVSIEDHREFDRGIDITRIILGAFRLN 113
OY 61 LOSNSLOGGSLTQOLIKLTYFSTSDQTSRKQAEMLAQLQEKATKQELITLYYINK 120
Db 114 LOSNSLOGGSLTQOLIKLTYFSTSDQTSRKQAEMLAQLQEKATKQELITLYYINK 173
OY 121 YVMSNGNGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPEAAQDRRLVL 180
Db 174 YVMSNGNGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPEAAQDRRLVL 233
OY 181 SEMKNGYISAEOYKAVANTPTDGLQSLKSASNPAYMDNYLKEVINQVEERTGYNLLT 240
Db 234 SEMKNGYISAEOYKAVANTPTDGLQSLKSASNPAYMDNYLKEVINQVEERTGYNLLT 293
OY 241 TGMDEVYTNVDEAOKHLMIDYNTDEYVAYPDELOVASTIVDSNGKYIAQLGARHOSN 300
Db 294 TGMDEVYTNVDEAOKHLMIDYNTDEYVAYPDELOVASTIVDSNGKYIAQLGARHOSN 353
OY 301 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYNGTNPVYNNDR 360
Db 354 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYNGTNPVYNNDR 413
OY 361 GYFGNITLQYALQOOSRNPAYETLKNVGLNRAKPTFLNGLGIDYPSIHNSNAISSTTSSD 420
Db 414 VYFGNITLQYALQOOSRNPAYETLKNVGLNRAKPTFLNGLGIDYPSIHNSNAISSTTSSD 473
OY 421 KKYGASSEKMAAAYAAAFANGGTYYKPMYIHKVYVSDGSEKESNVGTAKMETTAYMMTD 480
Db 474 KKYGASSEKMAAAYAAAFANGGTYYKPMYIHKVYVSDGSEKESNVGTAKMETTAYMMTD 533
```

```
OY 481 MKKTVLYTGTGRNAYLAMPQAGKTGTSNYTDELEIKHKNKGVAPEMFGYTRKYSM 540
Db 534 MKKTVLAGTGTGAGALPMLPAQAGKTGTSNYTDELEIKHKNKGVAPEMFGYTRKYSM 593
OY 541 AVMTCYSNRLLPVLVNGTLVAAKYVRSMWTYLSGSGNPEDNMIPEGGLRNGEFVKNNGAR 600
Db 554 AVMTCYSNRLLPVLVNGTLVAAKYVRSMWTYLSGSGNPEDNMIPEGGLRNGEFVKNNGAR 653
OY 601 STWNSPAPQOQPPSTRESSSSSDSSSTPSTNNSTTTNNPNNNQOOSNTTTPDOQONN 660
Db 654 STWNSPAPQOQPPSTRESSSSSDSSSTPSTNNSTTTNNPNNNQOOSNTTTPDOQONN 713
OY 661 POPAQP 666
Db 714 POPAQP 719
```

## RESULT 11

S28036  
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 8250) (fragment)

C:Species: Streptococcus pneumoniae

C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 26-May-2000

C:Accession: S28036

R:Martin, C.; Sibold, C.; Hakenbeck, R.

EMBO J. 11, 3831-3836, 1992

A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of

A:Reference number: S28031; MUID:93010977

A:Accession: S28036

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-608 <MAR>

A:Cross-references: EMBL:X67871; NID:q47415; PIDN:CAA48071.1; PID:q47416

C:Superfamily: penicillin-binding protein 1B

Query Match 80.9%, Score 2820; DB 2; Length 608;  
Best Local Similarity 96.8%, Pred. No. 1e-155;  
Matches 537; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

```
OY 1 KIYDKKNQIADLGSERRVNAQANDIPDLVKAIVSIEDHREFDRGIDITRIILGAFRLN 60
Db 54 KIFDSKNELIADLGSERRVNAQANDIPDLVKAIVSIEDHREFDRGIDITRIILGAFRLN 113
OY 61 LOSNSLOGGSLTQOLIKLTYFSTSDQTSRKQAEMLAQLQEKATKQELITLYYINK 120
Db 114 LOSNSLOGGSLTQOLIKLTYFSTSDQTSRKQAEMLAQLQEKATKQELITLYYINK 173
OY 121 YVMSNGNGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPEAAQDRRLVL 180
Db 174 YVMSNGNGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPEAAQDRRLVL 233
OY 181 SEMKNGYISAEOYKAVANTPTDGLQSLKSASNPAYMDNYLKEVINQVEERTGYNLLT 240
Db 234 SEMKNGYISAEOYKAVANTPTDGLQSLKSASNPAYMDNYLKEVINQVEERTGYNLLT 293
OY 241 TGMDEVYTNVDEAOKHLMIDYNTDEYVAYPDELOVASTIVDSNGKYIAQLGARHOSN 300
Db 294 TGMDEVYTNVDEAOKHLMIDYNTDEYVAYPDELOVASTIVDSNGKYIAQLGARHOSN 353
OY 301 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYNGTNPVYNNDR 360
Db 354 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYNGTNPVYNNDR 413
OY 361 GYFGNITLQYALQOOSRNPAYETLKNVGLNRAKPTFLNGLGIDYPSIHNSNAISSTTSSD 420
Db 414 GYFGNITLQYALQOOSRNPAYETLKNVGLNRAKPTFLNGLGIDYPSIHNSNAISSTTSSD 473
OY 421 KKYGASSEKMAAAYAAAFANGGTYYKPMYIHKVYVSDGSEKESNVGTAKMETTAYMMTD 480
Db 474 KKYGASSEKMAAAYAAAFANGGTYYKPMYIHKVYVSDGSEKESNVGTAKMETTAYMMTD 533
OY 481 MKKTVLYTGTGRNAYLAMPQAGKTGTSNYTDELEIKHKNKGVAPEMFGYTRKYSM 540
```



Db 534 MMKIVLVITIGIRGAYLWLPQAGTGTNTDEIEIKYINNTGYVADDEMFVGYTRKIAN 593

QY 541 AVMTGYSNRLEPLVG 555  
|||||  
594 AVMTGYSNRLEPLIG 608

Db

RESULT 12  
B42893

penicillin-binding protein 1 - Streptococcus oralis  
C:Species: Streptococcus oralis  
C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 04-Mar-2000  
C:Accession: B42893  
R:Martin, C./ Brisse, T.; Hakenbeck, R.  
J. Bacteriol. 174, 4517-4523, 1992  
A:Title: Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus oralis A and 1B.  
A:Reference number: A42893; MUID:92325042  
A:Accession: B42893  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-637 <MAR>  
A:Cross-references: GB:M0528; NID:g153769; PIDN:AAA26958.1; PID:g153771  
A:Superfamily: penicillin-binding protein 1B

Query Match	77.5%	Score 2698.5	DB 2	Length 637
Best Local Similarly	87.5%	Pred. No. 1.2e-146		
Matches 511	Conservative 39	Mismatches 33	Indels 1	Gaps 1

QY	1	KIINKNOLLADGSESRVNAQANDITDYLKAVSIEDHREFDNRGIDTIRILGAFLRN	60
Db	54	KIYNNDELLADGSESRVNAQAEIPLDYNALVSIEDHREFNRHRIIDTIRILGATLRN	113
QY	61	LO-SNSLQSGSTLQOLIKLTFYESTSDOTISSKAQEAMIALOLEBKAKFOELIYYIN	119
Db	114	LRRGGGLQAGASTLTQOLIKLTFYESTSDOTLSKKAQEAHLAVOLEBKAKFOELIYYIN	173
QY	120	KVYMSNGYNGMOTAAQNYCGKDLNLSL.POLALLAGMPQAPNQYDPYSHEPAAODRRNLV	179
Db	174	KVYMSNGYNGMOTAAQSYGCKDLKDLIS.POLALLAGMPQAPNQYDPYSHEPAAQERRNLV	233
QY	180	LSEKKNQGYISAEQYERKAVNTPITDGLQSLKSAASNRYAVNDYIKLEYINQVEETGYNL	239
Db	234	LSEKKGQGYITAEQYERKAIINTPITDGLQSLKSAANSRYPYNDYIKLEYIDQVEQETGYNL	293
QY	240	TTGMDVYTNVQDEQOKHLMIDYNTNDEVVAAPDDELQVASTIVYVSNCKVTAOLGARHOS	299
Db	294	TTGMEVYTNVQDSKQOQRKMDYNTNDEVVNPDEDLQVASTIVYVTDGKVTALQGARHOS	353
QY	300	NVSEGINQAVETNRDMGSTMKPITDYAPALEGYGYDSTATIVHDEPNYPGTNPVYNNMD	359
Db	354	NVSEGINQAVETNRDMGSTMKPITDYAPALEGYGYDSTASIVHDSPNYNGSTPYVNNMD	413
QY	360	RGYRGNTLLOYALQOOSRNPVAVETLKNVGLNRKATFLNGICIDYPSIHYSNALISSNTTES	419
Db	414	KSYEGNTLLOYALQOOSRNPVAVETLKVGLNRKATFLNGICIDYPSIHYANALISSNTTES	473
QY	420	DKKYGASEKKMAAAYAAPFANGTGYKRMVYIHKVYFSGSGSEKESNVGSTRAMKETTAVMNT	479
Db	474	DKKYGASEKKMAAAYAAPFANGTYPKPMYIINKLYFSDGSSKEFSDSGSTRAMKETTAVMNT	533
QY	480	DMKTTVLVLYGTGRNAYLAMLPOAQKCTSNVTDEEINHNKTSQFVAPDEDLFGAGYRRKS	539
Db	534	DMKTTVLVLYGTGRNAYLPMLPOAQKCTGNSVNTDEIENYIKNNGYVAPDMLFGYRRKS	593
QY	540	MAVNTGYSNRLTLPVYNGGLVYAAKAVYYSMAATYISEGNSPEDNMNI	583
Db	594	MAVNTGYSNRLTLPVYNGGFVYAAKAVYYSMAATYISEDDHNPDDNMNI	637

RESULT 13  
G866692

penicillin-binding protein 1A [imported] - Lactococcus lactis subsp. lactis (strain I  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
A:Accession: G86692  
R:Bohlin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarne, K.; Weissenbach, J.; Eh  
genome Res. 11, 731-753, 2001  
A:File: The complete genome sequence of the lactic acid bacterium Lactococcus lactis  
A:Reference number: A86625; MUID:21235186; PMID:11337471  
A:Accession: G86692  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-664 <STD>  
A:Cross-references: GI:AE005176; PID:912723428; PIDN:AAK04641.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: pona  
C:Superfamily: penicillin-binding protein 1B

Query Match	47.6%	Score 1657.5	DB 2	Length 664
Best Local Similarity	52.1%	Pred. No. 2.6e-88		
Matches 333	Conservative 105	Mismatches 166	Indels 35	Gaps 6

[illegible]

```

RESULT 14
AD1311
penicillin-binding protein 2A homolog pbpa [imported] - Listeria monocytogenes (strain
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

```

C:Accession: AD1311  
R:Glaser, P.; Frangenl, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fahl, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Stimes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1311  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1827 <GUA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC99970.1; PID:g16411345; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: pbpa

Query Match 34.9%; Score 1217; DB 2; Length 827;  
Best Local Similarity 40.3%; Pred. No. 1.2e-62;  
Matches 268; Conservative 133; Mismatches 210; Indels 54; Gaps 14;

1 KIYDNKQIADLGSERRVNAQNDIPTDYKAVSTEDHREFHGRIGITRIGAFLRN 60  
79 KLLDKDKVFAEVGTERREYIEKDIPELTKNALLTTEDEARFEDHDIDPRLGGAVIAN 138  
61 LOSN-SLOGSGSTTLOQIKITLFEFSTSDQTSRKAQOEWALIOLEKATKOELITYIN 119  
139 LTGQFGAGASTLSQOITIKSYL--DYTNKTLARKAOEWALIOLEKYSKNDILETYIN 196  
120 KYVMSNGNYMOTPAQNYKYKDLNNLSLPOLALLAGPQAPNOYDPYSHPEAADRNLV 179  
197 KYVMSDVRHGMQTAHEHYFEGKNVMDLTALQATALLAGPQSPNNPNPYEHPAAKKRRDY 256  
180 LSEKNGYISAQOYEKAVNPTITDGLQSLKASNPAYMDNYKEYINQVEETGNL 239  
257 LTNMYTHDKITKEEMTAQKPTITGTRSKKDRDKLYKDSYTVQVLSLEPKK--YD 314  
240 TTGMADVNTNDOEAKHLMIDYNTDEVAYAPDELOVASTIVDSNCKVIAOLGARHSS 299  
315 RDGLITHTALDRPAQETETKMLNTNFTDDEMAGIYLODKTRGVAIGGRNQ-- 372  
300 NVSEGINQAVETNRDMSYKPTITDPALEGYVSTATIVHDEPNYPGTNPVYND 359  
373 KYRGNVYATQVRSVSTKPIADYGPAPFEDLW--STAHLDEDEPYTSG--GTPINNMD 430  
360 RGEFNTTLOALQOOSNRNPAVEFLNKVGNRAKTFPLNGIGIDYPSIH--YSNAISSNT 417  
431 EGTGKPLSVQALYQSNIPALKTLOAVGLDKSEEFVNLGITYDEQONVESNAIGANS- 489  
418 ESDKKGASSEKMAAAYAFANGITYKPMYIKHVFSDD--SEKEFSNVGTTRAMKETAY 476  
490 -----SNMQAAGAAALGNGKGYNKPHYIKIYVSDQTEIDTPOSTVAMKESTAY 542  
477 MMTDMKTYVLTGTGRNAYLAMPQAGKTGTSNTDEIEINHKTISOFPV-----DEL 531  
543 MVSADVLTGISTGSAAPGVPAAGKTGTNIPE-----FTSKYYYPGSAARDSWF 596  
532 AGTTRKYSMAVMTGYSNRRLPLVNGILTVAAKYRSMYTLSEGSNPEDMNIBGLYRNG 591  
597 AGTYYNSTAVMTGYDCKKRYVASDQKIAQRNFSKIMAHVASAGKTADPKMPSNV-- 653  
592 EEFVKNGARSTWNSPAP-----QOPSTESSSSSDSSTQ--SSSTPSTNNSTTTNPN 645  
654 -----SVPIILKSNPIARAAGCTSSDKVSYELFLSGTAPTKTASTPEDEKK 699  
646 NMQOS 650  
700 KAEA 704  
RESULT 15

AD1683  
penicillin-binding protein 2A homolog pbpa [imported] - *Listeria innocua* (strain Clp  
C:Species: *Listeria innocua*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AD1683  
R:Glaser, P.; Frangenl, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fahl, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Stimes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1683  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1826 <GUA>  
A:Cross-references: GB:AL592022; PIDN:CAC97236.1; PID:g16414507; GSPDB:GN00178  
A:Experimental source: strain Clp11262  
C:Genetics:  
A:Gene: pbpa

Query Match 34.5%; Score 1201; DB 2; Length 826;  
Best Local Similarity 38.9%; Pred. No. 1e-61;  
Matches 264; Conservative 142; Mismatches 229; Indels 44; Gaps 14;

1 KIYDNKQIADLGSERRVNAQNDIPTDYKAVSTEDHREFHGRIGITRIGAFLRN 60  
79 KLLDKDKVFAEVGTERREYIEKDIPELTKNALLTTEDEARFEDHDIDPRLGGAVIAN 138  
61 LOSN-SLOGSGSTTLOQIKITLFEFSTSDQTSRKAQOEWALIOLEKATKOELITYIN 119  
139 VTGQFGAGASTLSQOITIKSYL--DYTNKTLARKAOEWALIOLEKYSKNDILETYIN 196  
120 KYVMSNGNYMOTPAQNYKYKDLNNLSLPOLALLAGPQAPNOYDPYSHPEAADRNLV 179  
197 KYVMSDVRHGMQTAHEHYFEGKNVMDLTALQATALLAGPQSPNNPNPYEHPAAKKRRDY 256  
180 LSEKNGYISAQOYEKAVNPTITDGLQSLKASNPAYMDNYKEYINQVEETGNL 239  
257 LTNMYTHDKITKEEMTAQKPTITGTRSKKDRDKLYKDSYTVQVLSLEPKK--YD 314  
240 TTGMADVNTNDOEAKHLMIDYNTDEVAYAPDELOVASTIVDSNCKVIAOLGARHSS 299  
315 RDGLITHTALDRPAQETETKMLNTNFTDDEMAGIYLODKTRGVAIGGRNQ-- 372  
300 NVSEGINQAVETNRDMSYKPTITDPALEGYVSTATIVHDEPNYPGTNPVYND 359  
373 KYRGNVYATQVRSVSTKPIADYGPAPFEDLW--STAHLDEDEPYTSG--GTPINNMD 430  
360 RGEFNTTLOALQOOSNRNPAVEFLNKVGNRAKTFPLNGIGIDYPSIH--YSNAISSNT 417  
431 EGTGKPLSVQALYQSNIPALKTLOAVGLDKSEEFVNLGITYDEQONVESNAIGANS- 489  
418 ESDKKGASSEKMAAAYAFANGITYKPMYIKHVFSDD--SEKERSNVGTTRAMKETAY 476  
490 -----SNMQAAGAAALGNGKGYNKPHYIKIYVSDQTEIDTPOSTVAMKESTAY 542  
477 MMTDMKTYVLTGTGRNAYLAMPQAGKTGTSNTDEIEINHKTISOFPV-----DEL 531  
543 MVSADVLTGISTGSAAPGVPAAGKTGTNIPE-----FTSKYYYPGSAARDSWF 596  
532 AGTTRKYSMAVMTGYSNRRLPLVNGILTVAAKYRSMYTLSEGSNPEDMNIBGLYRNG 591  
597 AGTYYNSTAVMTGYDCKKRYVASDQKIAQRNFSKIMAHVASAGKTADPKMPSNV--S 654  
592 EEFVKNGARSTWNSPAPQOPSTESSSS--SDSSTOSSSTTPTNNSTTTNPN 646  
655 VPIILG-----SNPIARAAGTASDKVSYELFLSGTAPTKTASTPEDEKKAEK 708  
647 TQOSNTTPOQONPOPAQ 665  
:: : : : :

Db 709 AEDKKTEEEKKEEPAK 727

Search completed: August 11, 2002, 04:24:15  
Job time: 4346 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 04:25:34 ; Search time 36.95 Seconds

(without alignments)  
697.895 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KIVDNKNQLADIGSERV.....TQSSNTTPDQGNQNPQAP 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3484	100.0	719	1	PPPA_STRPN
2	2698.5	77.5	637	1	PPPA_STROR
3	999	28.7	914	1	PPPA_BACSU
4	739	21.2	714	1	PPPE_BACSU
5	712	20.4	726	1	PPPA_AQUAE
6	680.5	19.5	798	1	PPPA_NEITL
7	642.5	18.4	822	1	PPPA_PSEAE
8	640	18.4	798	1	PPPA_NEIMA
9	640	18.4	798	1	PPPA_NEIMA
10	635.5	18.2	797	1	PPPA_NEIGO
11	633	18.2	798	1	PPPA_NEICI
12	617	17.7	825	1	PPPA_VIBCH
13	598.5	17.0	624	1	PPPD_BACSU
14	591.5	17.0	853	1	PPPA_HAEIN
15	584	16.8	777	1	PPPV_VIBCH
16	567	16.3	850	1	PPPA_ECOLI
17	526	15.1	844	1	PPPB_ECOLI
18	525.5	15.1	781	1	PPPB_HAEIN
19	502.5	14.4	787	1	PPPA_RICPR
20	490.5	14.1	809	1	PPPA_XYFPA
21	472	13.5	760	1	PPPB_BUCAI
22	450.5	12.9	770	1	PPPC_ECOLI
23	312.5	9.0	707	1	TRG_ALCEU
24	207	5.9	716	1	PPPB_BACSU
25	206	5.9	233	1	MTGA_NEIGO
26	206	5.9	645	1	SPSD_BACSU
27	194.5	5.6	224	1	MTGA_ACICA
28	193.5	5.6	598	1	FTSI_MESVI
29	192	5.5	233	1	MTGA_NEIMA
30	185.5	5.3	242	1	MTGA_KLEPM
31	185.5	5.3	242	1	MTGA_ECOLI
32	181	5.2	246	1	MTGA_HAEIN
33	164	4.7	588	1	FTSI_ECOLI

34	147	4.2	1861	1	APU_THETU	P36536 t amylopull
35	142.5	4.1	750	1	PPBX_STRPN	P14677 streptococ
36	142	4.1	610	1	FTSI_HAEIN	P45059 haemophilus
37	141	4.0	716	1	YOGF_BACSU	P54488 bacillus su
38	141	4.0	2843	1	APC_HUMAN	P25054 homo sapien
39	140.5	4.0	516	1	P54_ENTFC	P13692 enterococcu
40	139.5	4.0	706	1	PLB2_YEAST	Q03674 saccharomyc
41	138	4.0	1337	1	DEXT_STROO	P39653 streptococc
42	136.5	3.9	567	1	CHT3_CANAL	P40954 candida alb
43	136	3.9	579	1	FTSI_BUCAI	P57317 buchnera ap
44	135.5	3.9	1007	1	V741_CHLMG	Q9p16 chlamydia m
45	135	3.9	1365	1	GTF5_STRDO	P29336 streptococc

## ALIGNMENTS

RESULT ID	PPPA_STRPN	STANDARD:	PRT:	719 AA.
AC	004707			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	penicillin-binding protein 1A (PBP-1A) (Exported protein 2).			
GN	POMA OR EXP2 OR SP0369.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1313;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=45607, AND 63915;			
RX	MEDLINE=93010977; PubMed=1396576;			
RA	Martin C., Sibold C., Hakenbeck R.;			
RT	"Relatedness of penicillin-binding protein 1a genes from different			
RT	clones of penicillin-resistant Streptococcus pneumoniae isolated in			
RT	South Africa and Spain".;			
RL	EMBO J. 11:3831-3836(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=R6;			
RX	MEDLINE=92325042; PubMed=1624444;			
RA	Martin C., Briese T., Hakenbeck R.;			
RT	"Nucleotide sequences of genes encoding penicillin-binding proteins			
RT	from Streptococcus pneumoniae and Streptococcus oralis with high			
RT	homology to Escherichia coli penicillin-binding proteins 1a and 1b.";			
RL	J. Bacteriol. 174:4517-4523(1992).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=TRIG4;			
RX	MEDLINE=21357209; PubMed=11463916;			
RA	Tetelsen H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,			
RA	Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,			
RA	Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,			
RA	Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,			
RA	Holtzaple E., Knouri H., Wolf A.M., Ueberback T.R., Hansen C.L.,			
RA	McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,			
RA	Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,			
RA	Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;			
RT	"Complete genome sequence of a virulent isolate of Streptococcus			
RT	pneumoniae.";			
RL	Science 293:498-506(2001).			
RL	[4]			
RN	SEQUENCE OF 293-369 FROM N.A.			
RP	STRAIN=R6X;			
RC	MEDLINE=95020625; PubMed=7934910;			
RX	PEARCE B.J., Yin Y.B., Masure H.R.;			
RA	"Genetic identification of exported proteins in Streptococcus			
RT	pneumoniae.";			
RT	Mol. Microbiol. 9:1037-1050(1993).			
RL	-I- FUNCTION: CELL WALL FORMATION.			
CC	-I- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.			

```
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THESE STRAINS ARE PENICILLIN-SENSITIVE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X67873; CAA48073.1; -
DR EMBL: X67872; CAA48072.1; -
DR EMBL: M90527; AAA26956.1; -
DR EMBL: AE007349; AAK74536.1; -
DR PIR: S28038; S28038.
DR TIGR: SP0369; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
KW Multifunctional enzyme; Complete proteome.
FT ACCT_SITE 370 370 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT DOMAIN 658 683 SER-RICH.
FT VARIANT 124 124 T -> A (IN STRAIN R6).
FT VARIANT 386 386 V -> I (IN STRAIN 63915).
FT VARIANT 388 388 D -> E (IN STRAIN R6).
FT VARIANT 397 397 E -> K (IN STRAIN 63915).
FT VARIANT 523 523 M -> I (IN STRAIN 63915).
FT VARIANT 533 533 D -> E (IN STRAIN 45607).
FT VARIANT 540 540 T -> S (IN STRAINS 63915 AND R6).
FT VARIANT 657 657 T -> S (IN STRAINS 45607 AND R6).
SQ SEQUENCE 719 AA; 79758 MW; 5BD397E83B43BA6 CRC64;

Query Match 100.0%; Score 3484; DB 1; Length 719;
Best Local Similarity 100.0%; Pred. No. 2.5e-193;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KIYDNKNLIDLGSERVRNAQANDIPDLVKAIVSIEDHFRFDHGRGIDTIRILGAFLRN 60
DB 54 KIYDNKNLIDLGSERVRNAQANDIPDLVKAIVSIEDHFRFDHGRGIDTIRILGAFLRN 113
OY 61 LOSNSLOGSSTLTQOLIKLTYFSTSPDTISRKAQEWALAIQLEOKATKOEILTYINK 120
DB 114 LOSNSLOGSSTLTQOLIKLTYFSTSPDTISRKAQEWALAIQLEOKATKOEILTYINK 173
OY 121 YMSNGNNGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNQYDPSHPPEAADRRLVL 180
DB 174 YMSNGNNGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNQYDPSHPPEAADRRLVL 233
OY 181 SEMKNOGYSAEQYKAVNTPTDGLQSLKSASNPATMDNTLKVYNQVEETGYNLIT 240
DB 234 SEMKNOGYSAEQYKAVNTPTDGLQSLKSASNPATMDNTLKVYNQVEETGYNLIT 293
OY 241 TGMDEVTVNDEQAKHLDIYNTDEYVAYPPDELOVASTYDVNSGKYIAOLGARRHOSN 300
DB 294 TGMDEVTVNDEQAKHLDIYNTDEYVAYPPDELOVASTYDVNSGKYIAOLGARRHOSN 353
OY 301 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGNTPTVYNNDR 360
DB 354 VSFGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGNTPTVYNNDR 413
OY 361 GYFGNITITQYALQOSRNPAYETLTKVGLNRAKTFNLGLIDYPSIHSMASITSTSD 420
DB 414 GYFGNITITQYALQOSRNPAYETLTKVGLNRAKTFNLGLIDYPSIHSMASITSTSD 473
OY 421 KKYGASSEKMAAAYAFANGGTYYKPYTHKVVPSDSEKESNNGTRAMKETTYMMTD 480
DB 474 KKYGASSEKMAAAYAFANGGTYYKPYTHKVVPSDSEKESNNGTRAMKETTYMMTD 533
```

```
OY 481 MKKVLVTGTGNAYLAWLPQAGKTGTSNYTDEEIEHNHKTISOVAPDELFAGYTRKYSM 540
DB 534 MKKVLVTGTGNAYLAWLPQAGKTGTSNYTDEEIEHNHKTISOVAPDELFAGYTRKYSM 593
OY 541 AWTGTYSNLTPTLVNGGLTVAAKVYRSMWYTLSECSNPEDMWIIPGIVRNGEYFKNAR 600
DB 594 AWTGTYSNLTPTLVNGGLTVAAKVYRSMWYTLSECSNPEDMWIIPGIVRNGEYFKNAR 653
OY 601 STWNSPAPQPPSTSSSSSDSTSSSTPTPTNNSTTTPNNNTPOQSNTPPOQON 660
DB 654 STWNSPAPQPPSTSSSSSDSTSSSTPTPTNNSTTTPNNNTPOQSNTPPOQON 713
OY 661 PQAPQ 666
DB 714 PQAPQ 719

RESULT 2
PBPA_STROR STANDARD; PRT; 637 AA.
ID PBPA_STROR
AC 000573;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Penicillin-binding protein 1A (PBP-1A) (Fragment).
GN PONA.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus;
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92325042; Pubmed=1624444;
RA Martin C., Biese T., Hakenbeck R.;
RT "Nucleotide sequences of genes encoding penicillin-binding proteins
RT from Streptococcus pneumoniae and Streptococcus oralis with high
RT homology to Escherichia coli penicillin-binding proteins 1a and 1b.";
RL J. Bacteriol. 174:4517-4523(1992).
CC -1- FUNCTION: CELL WALL FORMATION.
CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M90528; AAA26958.1; -
DR PIR: B42893; B42893.
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
KW Multifunctional enzyme.
FT ACCT_SITE 371 371 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT NON_TER 637 637
SQ SEQUENCE 637 AA; 70891 MW; A6D198BCEA603A63 CRC64;

Query Match 77.5%; Score 2698.5; DB 1; Length 637;
Best Local Similarity 87.5%; Pred. No. 3.5e-148;
Matches 511; Conservative 39; Mismatches 33; Indels 1; Gaps 1;

OY 1 KIYDNKNLIDLGSERVRNAQANDIPDLVKAIVSIEDHFRFDHGRGIDTIRILGAFLRN 60
DB 54 KIYDNKNLIDLGSERVRNAQANDIPDLVKAIVSIEDHFRFDHGRGIDTIRILGAFLRN 113
OY 61 LOSNSLOGSSTLTQOLIKLTYFSTSPDTISRKAQEWALAIQLEOKATKOEILTYINK 119
```

```

114 LRGGGLGASTLTQOOLIKLTFSTSTSDPTLSRKAQEMLAVALQLEQKATKQELLTYIN 173
120 KYMNSNGNGMGTAFANONYGRDLNLSLPQALLAGMPQAPQOPYPYSHPEAAQDRNLV 179
124 KYMNSNGNGMGTAFANONYGRDLNLSLPQALLAGMPQAPQOPYPYSHPEAAQDRNLV 233
180 LSEMNQGYISAEQYERKAVNPTITDGLSLKASNNYPAMNDYKLEKVINQVEEENGYNL 239
234 LSEMGQGYITAEQYERKAVNPTITDGLSLKASNNYPAMNDYKLEKVINQVEEENGYNL 293
240 TTGMDEVYTNVDOEAKRHLMDIYNTDEYVAYPPDELOVASTIYVSNKGVIQALGARHSS 299
294 TTGMDEVYTNVDOEAKRHLMDIYNTDEYVAYPPDELOVASTIYVSNKGVIQALGARHSS 353
300 NVSEGINQAVEYENRDMGSMKPTIDYAPALEGYVDSTATYHDEPYNPGTINTPYNMD 359
334 NVSEGINQAVEYENRDMGSMKPTIDYAPALEGYVDSTATYHDEPYNPGTINTPYNMD 413
360 RGYFGNITLOVALQOOSRNPVAVETLNKVLNRAKTFPLNGLGIDYPSIHYSNMISSTTES 419
414 KYPFGNITLOVALQOOSRNPVAVETLNKVLNRAKTFPLNGLGIDYPSIHYSNMISSTTES 473
420 DKRYGASSEKMAAAYAAFAFANGGTYYKPMYIHKVYFSDGSEKESVNGTRAMKETTAYMT 479
474 DKRYGASSEKMAAAYAAFAFANGGTYYKPMYIHKVYFSDGSEKESVNGTRAMKETTAYMT 533
480 DMKATVLYTGTRNAYLAFLPAQAGTGTSTYDELEHNIKTQYAPDELFRAGTRKKS 539
534 DMKATVLYTGTRNAYLAFLPAQAGTGTSTYDELEHNIKTQYAPDELFRAGTRKKS 593
540 MAVWTGYSNRLTFLVNGNGLTVAAKYRSMYMTLSEGSNEDPMNI 583
594 MAVWTGYSNRLTFLVNGNGLTVAAKYRSMYMTLSEGSNEDPMNI 637

```

```

RA Murray T., Popham D.L., Setlow P.;
RT "Bacillus subtilis cells lacking penicillin-binding protein 1 require
RT increased levels of divalent cations for growth."
RL J. Bacteriol. 180:4555-4563(1998).
RN [4]
RP SUBCELLULAR LOCATION.
RC STRAIN-168;
RX MEDLINE-99255546; PubMed-10322023;
RA Pedersen L.B., Angert E.R., Setlow P.;
RT "Septal localization of penicillin-binding protein 1 in Bacillus
RT subtilis."
RL J. Bacteriol. 181:3201-3211(1999).
CC -I- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC SUBUNITS) (BY SIMILARITY).
CC -I- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -I- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. PROBABLY FOUND ALL
CC OVER THE WHOLE CELL AT LOW CONCENTRATIONS. ALSO LOCALIZES TO THE
CC DIVISION SITE IN VEGETATIVE CELLS.
CC -I- DEVELOPMENTAL STAGE: EXPRESSION IS CONSTANT DURING GROWTH,
CC DECREASES DURING SPOROULATION AND IS INDUCED APPROXIMATELY 15 MIN
CC INTO SPORE GERMINATION.
CC -I- PTM: THE PRODUCT EXPRESSED FROM THE TRANSLATION OF THE PONA GENE
CC APPEARS AS TWO BANDS ON A GEL (1A AND 1B), BUT THE SPECIFIC AMINO
CC ACID SEQUENCE OF EACH PROTEIN IS UNKNOWN.
CC -I- PTM: THE N-TERMINUS IS BLOCKED.
CC -I- MISCELLANEOUS: CELLS LACKING THE PROTEIN REQUIRE INCREASED LEVELS
CC OF MG2+ OR CA2+ FOR GROWTH AND GERMINATION. APPROXIMATELY 50% OF
CC CELLS WITHOUT THE PROTEIN CONTAIN ABNORMAL ATP2 RINGS, SUGGESTING
CC IT IS INVOLVED IN SEPTUM SYNTHESIS; INCREASED LEVELS OF MG2+ OR
CC CA2+ ONLY PARTIALLY ELIMINATE THE SEPTATION DEFECTS.
CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC TRANSGLYCOSYLASE FAMILY.
CC -I- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRANSPEPTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, U11883; AAA64947.1; -
DR EMBL, L47838; AAB38459.1; -
DR EMBL, Z99115; CAB14148.1; -
DR HSP, P02751; LFNA.
DR Subtilist; Bg10954; pona.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00041; fn3; 1.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
DR SMART; SM00060; FN3; 1.
DR Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
KW Hydrolyase; Multifunctional enzyme; Transmembrane; Signal-anchor;
KW Antibiotic resistance; Complete proteome.
KW DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
KW DOMAIN 38 58 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
KW TRANSMEM 1 37
KW TRANSMEM 38 58
FT DOMAIN 59 914 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 77 246 TRANSGLYCOSYLASE.
FT DOMAIN 329 662 TRANSPEPTIDASE.
FT ACT_SITE 390 390 ACTYLATED BY PENICILLIN (BY SIMILARITY).
SQ SEQUENCE 914 AA; 99562 MW; 697BE33DFE2423E6 CRC64;

```





Query Match 21.2%; Score 739; DB 1; Length 714;  
 Best Local Similarity 32.2%; Pred. No. 2,7e-35;  
 Matches 187; Conservative 119; Mismatches 237; Indels 38; Gaps 18;

QY 1 KIYDNKNQIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRIDITRILGAFARN 60  
 DB 50 KIYDQNGDEVASLYTEREPVSINEIPKQVREAFIATEDREFEHNGIDAKSVGRAYRD 109  
 QY 61 -LQSNLSGGSTLTLQOLIKLFYFSTSDOTISRKQAEMLATOLEKAKKOEITLYYIN 119  
 DB 110 ILAGKVEGGTTTQOLAKNIFL--THDKTELRKTEVITAINLERDYSKDLLEYLN 166  
 QY 120 KYVMSNGYGMQTAONYKDKLNLNLSLPOLALLAGMPAPNOVDPSHPEAADRNLV 179  
 DB 167 QLFEGHGVYIGQASHYIFPKKEVKDLTVSGAVLAIPKAPSTYSPILHDKKERDIT 226  
 QY 180 LSEKKNQYISAEQYEAQVAVTPTDGLQSLKSASNPAYMDNLYKEVINQVEE--TGY 236  
 DB 227 LGMNDQGYISAKEAVTAQGR--TLGLH-VKQSETP-WPDSYIDLIVKAEDEKYSISGE 282  
 QY 237 NLTGMDVYTNVDOEQKHLMDIYNTDEVVAIPDEDELQVASTIVDYSNCKVIAOLGARH 296  
 DB 283 QLLQGGTTIKVPLDSKLOKTAQVMKEGSYYPETDQNAESGSAVFINKTKGVEALIGRD 342  
 QY 297 OSSNVSFGINQAVETNRDMGSTMKPIITDYAPALEGYVDSTATIVHDEPNYPGTNTPVY 356  
 DB 343 YTKR---GYNR-VTAVRQPSSTFKPLAVYGPMAQEKKF-KPYSLKDELQSY-GDITP-K 395  
 QY 357 NMDRGYFNTITQYALQOSNRNPAVETLANKVGNRAKFTNGLIDVPSIHYNSAISNT 416  
 DB 396 NYDSREGEVYTMADAITYSKNAPAVWTLNEIGVETGSKYLKANGCIDIPDEGLALLGCG- 453  
 QY 417 TESDKRYGASSEMAAAYAFANGCITYKPMYIHKVYFSGSGE-KEPSNVTGAMKETTA 475  
 DB 454 ---LEKGVSPDLAGAFHTEFAANGYTEPEPFISIIDEGETIADHKEGKRVFSKOTS 509  
 QY 476 YWMTDMKVTLYTGTGRN-AYLAMLPQAGKTGTSNTDEIENHIKTSQVAPDELPAGY 534  
 DB 510 WNNTRMLQOVYKKGITATSGTYHGDLL-AGTGSTSTYTG--VSGATK-----DAMFAGY 558  
 QY 535 TRKYSMAVWTGYSNR-LTPLVNGCLTVAAKVYRSMYTLSE 574  
 DB 559 TPKITGAVWVGDKTDQNNHYLKAGSSYPTRLFDILTLQAGE 599

RESULT 5  
 ID BPBA\_AQDAE STANDARD; PRT; 726 AA.  
 AC 066874;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Penicillin-binding protein 1A (PBP-1a) (BPPIa) [Includes: Penicillin-  
 DE insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan TGase);  
 DE penicillin-sensitive transpeptidase (EC 3.4.4.-) (DD-transpeptidase)].  
 GN MRCA OR POMA OR AQ\_624.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 OX NCBI\_TaxID=63363;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RX MEDLINE=98196666; PubMed=9537320;  
 RA Decker G., Warren P.V., Gaasterland T., Young W.G., Tenox A.L.,  
 RA Grelham D.E., Overbeek R., Sneed M.A., Keller M., Autay M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 aeolicus".  
 RL Nature 392:353-358(1998).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

QY 1 KIYDNKNQIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRIDITRILGAFARN 60  
 DB 46 EYVDARKRLVGTIGIQKRFVYSIDKIPYEHVIMFVAREDRNFHMFGLDPAIVRAIAYN 105  
 QY 61 LQSNLS-OGSSTLTLQOLIKLFYFSTSDOTISRKQAEMLATOLEKAKKOEITLYYIN 119  
 DB 106 YRAGRIVQGGSTTQOLAKNIFL--TRERTLEKRIKIALTAIKIERTFPKKIMELYN 162  
 QY 120 KYVMSNGYGMQTAONYKDKLNLNLSLPOLALLAGMPAPNOVDPSHPEAADRNLV 179  
 DB 163 QYILGSAIVGEAAQVYFGHWHWELSLDEAALALAPKAPAKNPYHBERALQRRNLV 222  
 QY 180 LSEKKNQYISAEQYEAQVAVTPTDGLQSLKSASNPAYMDNLYKEVINQVEEYGNL 239  
 DB 223 LKRRLEGGYITPQGEAEAVKPLT---VKREKY-KFDGYFLDMVKYVFNKYGELAY 276  
 QY 240 TTGMDVYTNVDOE---AQHLMD----- 259  
 DB 277 KGRILKIYTTIDLYQKIAQKSLGELKRVAKIIGLPKPSSEDMELAYEKEAOLKRLKR 336  
 QY 260 --TY-----PDDDELQY 276  
 DB 337 GKTYVAKILYDGNFMKVEIHGKKLKEIGLNTGKHVYFVKYLGNNRAELIIPDLGSL 396  
 QY 277 ASTIVDYSNGKVIQAOLGARQSSNVSFGINQAVETNRDMGSTMKPIITDYAPALEGYVD 336  
 DB 397 VS-IDVKTEIKAIIVGR---SVAYSQFNRVAKALNPOPSAIKPYI-YLSALLKGMTQI 450  
 QY 337 TATIVHDEPYNYP-GTNTPVYKMDRGYFGNITLOVALQOSNRNPAVETLANKVGNRAKT 394  
 DB 451 STIDASSKPYDPSKGEDMIPKNYDEKEYGNVTLRYALASHIMTAVALNLDKGFGFELVLE 510  
 QY 395 FLNGLIGDYSIHYSNASSNTTESDKKYGASSKMAAAYAFANGSTYKPMYIHKVYF 454  
 DB 511 VKKKVGIDNLKPYISLAL-GTVE-----VTPQLTLAAOVFANLGTCEKPFPIKRIYD 562  
 QY 455 SDSEKEFFSNVG--TRAKKETTAVMYMDMKTYLTYTGCTGNAYLAWLPQAGKTGTSNYTD 512  
 DB 563 ENGEVLE-ENVPBEQVLELPRPRTVYDMLRAVLBETARASVLDIVAGKGT---FD 618  
 QY 513 EETENHIKTSQFVAPDELPAELGAYTRKYSMAVWGYSNLPLVVG---NGLTVAAKVYSKM 569  
 DB 619 D-----FODAMFVGFSPYIVTGVWVGVDYKKS--LGKHMGSRYALPITWIDM 664  
 QY 570 TYLSEGSNPEDWNIP 584

Query Match 20.4%; Score 712; DB 1; Length 726;  
 Best Local Similarity 30.7%; Pred. No. 1e-33;  
 Matches 207; Conservative 109; Mismatches 227; Indels 132; Gaps 21;

QY 1 KIYDNKNQIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRIDITRILGAFARN 60  
 DB 46 EYVDARKRLVGTIGIQKRFVYSIDKIPYEHVIMFVAREDRNFHMFGLDPAIVRAIAYN 105  
 QY 61 LQSNLS-OGSSTLTLQOLIKLFYFSTSDOTISRKQAEMLATOLEKAKKOEITLYYIN 119  
 DB 106 YRAGRIVQGGSTTQOLAKNIFL--TRERTLEKRIKIALTAIKIERTFPKKIMELYN 162  
 QY 120 KYVMSNGYGMQTAONYKDKLNLNLSLPOLALLAGMPAPNOVDPSHPEAADRNLV 179  
 DB 163 QYILGSAIVGEAAQVYFGHWHWELSLDEAALALAPKAPAKNPYHBERALQRRNLV 222  
 QY 180 LSEKKNQYISAEQYEAQVAVTPTDGLQSLKSASNPAYMDNLYKEVINQVEEYGNL 239  
 DB 223 LKRRLEGGYITPQGEAEAVKPLT---VKREKY-KFDGYFLDMVKYVFNKYGELAY 276  
 QY 240 TTGMDVYTNVDOE---AQHLMD----- 259  
 DB 277 KGRILKIYTTIDLYQKIAQKSLGELKRVAKIIGLPKPSSEDMELAYEKEAOLKRLKR 336  
 QY 260 --TY-----PDDDELQY 276  
 DB 337 GKTYVAKILYDGNFMKVEIHGKKLKEIGLNTGKHVYFVKYLGNNRAELIIPDLGSL 396  
 QY 277 ASTIVDYSNGKVIQAOLGARQSSNVSFGINQAVETNRDMGSTMKPIITDYAPALEGYVD 336  
 DB 397 VS-IDVKTEIKAIIVGR---SVAYSQFNRVAKALNPOPSAIKPYI-YLSALLKGMTQI 450  
 QY 337 TATIVHDEPYNYP-GTNTPVYKMDRGYFGNITLOVALQOSNRNPAVETLANKVGNRAKT 394  
 DB 451 STIDASSKPYDPSKGEDMIPKNYDEKEYGNVTLRYALASHIMTAVALNLDKGFGFELVLE 510  
 QY 395 FLNGLIGDYSIHYSNASSNTTESDKKYGASSKMAAAYAFANGSTYKPMYIHKVYF 454  
 DB 511 VKKKVGIDNLKPYISLAL-GTVE-----VTPQLTLAAOVFANLGTCEKPFPIKRIYD 562  
 QY 455 SDSEKEFFSNVG--TRAKKETTAVMYMDMKTYLTYTGCTGNAYLAWLPQAGKTGTSNYTD 512  
 DB 563 ENGEVLE-ENVPBEQVLELPRPRTVYDMLRAVLBETARASVLDIVAGKGT---FD 618  
 QY 513 EETENHIKTSQFVAPDELPAELGAYTRKYSMAVWGYSNLPLVVG---NGLTVAAKVYSKM 569  
 DB 619 D-----FODAMFVGFSPYIVTGVWVGVDYKKS--LGKHMGSRYALPITWIDM 664  
 QY 570 TYLSEGSNPEDWNIP 584

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

QY 1 KIYDNKNQIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRIDITRILGAFARN 60  
 DB 46 EYVDARKRLVGTIGIQKRFVYSIDKIPYEHVIMFVAREDRNFHMFGLDPAIVRAIAYN 105  
 QY 61 LQSNLS-OGSSTLTLQOLIKLFYFSTSDOTISRKQAEMLATOLEKAKKOEITLYYIN 119  
 DB 106 YRAGRIVQGGSTTQOLAKNIFL--TRERTLEKRIKIALTAIKIERTFPKKIMELYN 162  
 QY 120 KYVMSNGYGMQTAONYKDKLNLNLSLPOLALLAGMPAPNOVDPSHPEAADRNLV 179  
 DB 163 QYILGSAIVGEAAQVYFGHWHWELSLDEAALALAPKAPAKNPYHBERALQRRNLV 222  
 QY 180 LSEKKNQYISAEQYEAQVAVTPTDGLQSLKSASNPAYMDNLYKEVINQVEEYGNL 239  
 DB 223 LKRRLEGGYITPQGEAEAVKPLT---VKREKY-KFDGYFLDMVKYVFNKYGELAY 276  
 QY 240 TTGMDVYTNVDOE---AQHLMD----- 259  
 DB 277 KGRILKIYTTIDLYQKIAQKSLGELKRVAKIIGLPKPSSEDMELAYEKEAOLKRLKR 336  
 QY 260 --TY-----PDDDELQY 276  
 DB 337 GKTYVAKILYDGNFMKVEIHGKKLKEIGLNTGKHVYFVKYLGNNRAELIIPDLGSL 396  
 QY 277 ASTIVDYSNGKVIQAOLGARQSSNVSFGINQAVETNRDMGSTMKPIITDYAPALEGYVD 336  
 DB 397 VS-IDVKTEIKAIIVGR---SVAYSQFNRVAKALNPOPSAIKPYI-YLSALLKGMTQI 450  
 QY 337 TATIVHDEPYNYP-GTNTPVYKMDRGYFGNITLOVALQOSNRNPAVETLANKVGNRAKT 394  
 DB 451 STIDASSKPYDPSKGEDMIPKNYDEKEYGNVTLRYALASHIMTAVALNLDKGFGFELVLE 510  
 QY 395 FLNGLIGDYSIHYSNASSNTTESDKKYGASSKMAAAYAFANGSTYKPMYIHKVYF 454  
 DB 511 VKKKVGIDNLKPYISLAL-GTVE-----VTPQLTLAAOVFANLGTCEKPFPIKRIYD 562  
 QY 455 SDSEKEFFSNVG--TRAKKETTAVMYMDMKTYLTYTGCTGNAYLAWLPQAGKTGTSNYTD 512  
 DB 563 ENGEVLE-ENVPBEQVLELPRPRTVYDMLRAVLBETARASVLDIVAGKGT---FD 618  
 QY 513 EETENHIKTSQFVAPDELPAELGAYTRKYSMAVWGYSNLPLVVG---NGLTVAAKVYSKM 569  
 DB 619 D-----FODAMFVGFSPYIVTGVWVGVDYKKS--LGKHMGSRYALPITWIDM 664  
 QY 570 TYLSEGSNPEDWNIP 584

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

QY 1 KIYDNKNQIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRIDITRILGAFARN 60  
 DB 46 EYVDARKRLVGTIGIQKRFVYSIDKIPYEHVIMFVAREDRNFHMFGLDPAIVRAIAYN 105  
 QY 61 LQSNLS-OGSSTLTLQOLIKLFYFSTSDOTISRKQAEMLATOLEKAKKOEITLYYIN 119  
 DB 106 YRAGRIVQGGSTTQOLAKNIFL--TRERTLEKRIKIALTAIKIERTFPKKIMELYN 162  
 QY 120 KYVMSNGYGMQTAONYKDKLNLNLSLPOLALLAGMPAPNOVDPSHPEAADRNLV 179  
 DB 163 QYILGSAIVGEAAQVYFGHWHWELSLDEAALALAPKAPAKNPYHBERALQRRNLV 222  
 QY 180 LSEKKNQYISAEQYEAQVAVTPTDGLQSLKSASNPAYMDNLYKEVINQVEEYGNL 239  
 DB 223 LKRRLEGGYITPQGEAEAVKPLT---VKREKY-KFDGYFLDMVKYVFNKYGELAY 276  
 QY 240 TTGMDVYTNVDOE---AQHLMD----- 259  
 DB 277 KGRILKIYTTIDLYQKIAQKSLGELKRVAKIIGLPKPSSEDMELAYEKEAOLKRLKR 336  
 QY 260 --TY-----PDDDELQY 276  
 DB 337 GKTYVAKILYDGNFMKVEIHGKKLKEIGLNTGKHVYFVKYLGNNRAELIIPDLGSL 396  
 QY 277 ASTIVDYSNGKVIQAOLGARQSSNVSFGINQAVETNRDMGSTMKPIITDYAPALEGYVD 336  
 DB 397 VS-IDVKTEIKAIIVGR---SVAYSQFNRVAKALNPOPSAIKPYI-YLSALLKGMTQI 450  
 QY 337 TATIVHDEPYNYP-GTNTPVYKMDRGYFGNITLOVALQOSNRNPAVETLANKVGNRAKT 394  
 DB 451 STIDASSKPYDPSKGEDMIPKNYDEKEYGNVTLRYALASHIMTAVALNLDKGFGFELVLE 510  
 QY 395 FLNGLIGDYSIHYSNASSNTTESDKKYGASSKMAAAYAFANGSTYKPMYIHKVYF 454  
 DB 511 VKKKVGIDNLKPYISLAL-GTVE-----VTPQLTLAAOVFANLGTCEKPFPIKRIYD 562  
 QY 455 SDSEKEFFSNVG--TRAKKETTAVMYMDMKTYLTYTGCTGNAYLAWLPQAGKTGTSNYTD 512  
 DB 563 ENGEVLE-ENVPBEQVLELPRPRTVYDMLRAVLBETARASVLDIVAGKGT---FD 618  
 QY 513 EETENHIKTSQFVAPDELPAELGAYTRKYSMAVWGYSNLPLVVG---NGLTVAAKVYSKM 569  
 DB 619 D-----FODAMFVGFSPYIVTGVWVGVDYKKS--LGKHMGSRYALPITWIDM 664  
 QY 570 TYLSEGSNPEDWNIP 584

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

QY 1 KIYDNKNQIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRIDITRILGAFARN 60  
 DB 46 EYVDARKRLVGTIGIQKRFVYSIDKIPYEHVIMFVAREDRNFHMFGLDPAIVRAIAYN 105  
 QY 61 LQSNLS-OGSSTLTLQOLIKLFYFSTSDOTISRKQAEMLATOLEKAKKOEITLYYIN 119  
 DB 106 YRAGRIVQGGSTTQOLAKNIFL--TRERTLEKRIKIALTAIKIERTFPKKIMELYN 162  
 QY 120 KYVMSNGYGMQTAONYKDKLNLNLSLPOLALLAGMPAPNOVDPSHPEAADRNLV 179  
 DB 163 QYILGSAIVGEAAQVYFGHWHWELSLDEAALALAPKAPAKNPYHBERALQRRNLV 222  
 QY 180 LSEKKNQYISAEQYEAQVAVTPTDGLQSLKSASNPAYMDNLYKEVINQVEEYGNL 239  
 DB 223 LKRRLEGGYITPQGEAEAVKPLT---VKREKY-KFDGYFLDMVKYVFNKYGELAY 276  
 QY 240 TTGMDVYTNVDOE---AQHLMD----- 259  
 DB 277 KGRILKIYTTIDLYQKIAQKSLGELKRVAKIIGLPKPSSEDMELAYEKEAOLKRLKR 336  
 QY 260 --TY-----PDDDELQY 276  
 DB 337 GKTYVAKILYDGNFMKVEIHGKKLKEIGLNTGKHVYFVKYLGNNRAELIIPDLGSL 396  
 QY 277 ASTIVDYSNGKVIQAOLGARQSSNVSFGINQAVETNRDMGSTMKPIITDYAPALEGYVD 336  
 DB 397 VS-IDVKTEIKAIIVGR---SVAYSQFNRVAKALNPOPSAIKPYI-YLSALLKGMTQI 450  
 QY 337 TATIVHDEPYNYP-GTNTPVYKMDRGYFGNITLOVALQOSNRNPAVETLANKVGNRAKT 394  
 DB 451 STIDASSKPYDPSKGEDMIPKNYDEKEYGNVTLRYALASHIMTAVALNLDKGFGFELVLE 510  
 QY 395 FLNGLIGDYSIHYSNASSNTTESDKKYGASSKMAAAYAFANGSTYKPMYIHKVYF 454  
 DB 511 VKKKVGIDNLKPYISLAL-GTVE-----VTPQLTLAAOVFANLGTCEKPFPIKRIYD 562  
 QY 455 SDSEKEFFSNVG--TRAKKETTAVMYMDMKTYLTYTGCTGNAYLAWLPQAGKTGTSNYTD 512  
 DB 563 ENGEVLE-ENVPBEQVLELPRPRTVYDMLRAVLBETARASVLDIVAGKGT---FD 618  
 QY 513 EETENHIKTSQFVAPDELPAELGAYTRKYSMAVWGYSNLPLVVG---NGLTVAAKVYSKM 569  
 DB 619 D-----FODAMFVGFSPYIVTGVWVGVDYKKS--LGKHMGSRYALPITWIDM 664  
 QY 570 TYLSEGSNPEDWNIP 584

DB 665 KVTATMPNDEFELP 679

RESULT 6

PPPA\_NEIFL STANDARD; PRT; 798 AA.

ID PPBA\_NEIFL

AC 087626;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Penicillin-binding protein 1A (PBP-1a) (PBP1a) [Includes: Penicillin-insensitive transglycosylase (EC 2.4.2.-) (peptidoglycan TGase)];

DE Penicillin-sensitive transpeptidase (EC 3.4.-.-) (Dd-transpeptidase)].

GN MRCA OR PONA.

OS Neisseria flavescens.

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

OX NCBI\_TaxId=484;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-NRL 30009;

RA Ropp P.A., Nicholas R.A.;

RT "Nucleotide sequence of the ponA gene encoding penicillin-binding protein 1 of Neisseria flavescens.";

RL Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE SUBUNITS) (BY SIMILARITY).

CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.

CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By similarity).

CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE TRANSGLYCOSYLASE FAMILY.

CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE TRANSPEPTIDASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Its usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: AF087677; AAC35856.1; .

DR InterPro: IPR001264; Transglycosyl.

DR InterPro: IPR001460; transpeptidase.

DR Pfam: PF00912; Transglycosyl; 1.

DR Pfam: PF00905; Transpeptidase; 1.

DR Prodom: PD001895; Transglycosyl; 1.

KM Peptidoglycan synthesis; Cell wall; Transferrase; Glycosyltransferase; Hydrolyase; Multifunctional enzyme; Transmembrane; Inner membrane; Signal-anchor; Antibiotic resistance.

KM Hydrolyase; Multifunctional enzyme; Transmembrane; Inner membrane.

FT TRANSMEM 1 9 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN) (POTENTIAL).

FT TRANSMEM 31 798 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 50 218 TRANSGLYCOSYLASE.

FT TRANSMEM 413 699 TRANSPEPTIDASE.

FT TRANSMEM 460 460 ACYLATED BY PENICILLIN (BY SIMILARITY).

SEQUENCE 798 AA: 87703 MW; 0DDCDF6FD5953AA CRC64;

Query Match 19.5%; Score 680.5; DB 1; Length 798; Best Local Similarity 25.8%; Pred. No. 7.5e-32; Matches 203; Conservative 129; Mismatches 277; Indels 179; Gaps 20;

QY 2 IYDNKNQIADIGSERVNAQNDIPTDLYKAIVSIEDHREFDHGIDITIRILGAFRLN 61

DB 52 IYSSGQVIGYGEQRREFTKIDPEKILKDAVIAEDKRFYDHGVDWGVARAIVGNV 111

QY 62 OSNSLQ-GGSLTLQOLIKLIVESTSTSDQTSISKRAEAMLAIOLEOKATKOELLTYINK 120

DB 112 MAGGVGSAGSTIQYAKNPFYLS---SERSFTFKFNEALLATYIEQSLSKDKLTLEYFNG 168

QY 121 VYWSNGNYGMOTAAONYGYKDLNNLSLPOLALLAGMPQAPNOYDPSHPBAODRNLV 180

DB 169 IYLGARAYGFASAAQTYEFKNVNDLTLAEAMLAGLPKAPSAVNPVNPBRALRQAYLL 228

QY 181 SEKKNGCYISABEYKAQVMTPTIDGLOSLSKASNYAVYNDN-----YKKEYNO-VVEE 233

DB 229 NNMLEGKMTTLQORD-----QALKEELHYERFVQNDISALYVEMARQELFER 277

QY 234 TGYNLTTCMDVT-----

DB 278 YGEDATYQGFKYTTTVDTAHQVATEALKVLRNPDGSSRYGAENVYDLSKSDNVEEV 337

QY 254 QKHLDIYNTDEVY-----

DB 338 SOYLSTLYTVDKMIPAVVLEASRKGYQIOLPSGRKYTLNNHALGFARAANNEKMGDDRI 397

QY 275 QVASTI-----VDVSNKGVIQOLGARHQSNVSGINQAVET 311

DB 398 RRGSVIRVKGSGDTFTVVOEPLLOGALVSLDAKTGAVRALVGIDYHSKT---FNRATQA 454

QY 312 NRDMGSTMPTIDYAPALEGYVDSTATTIHDEPNYPGTNTPVYNW-----DRGYFGNI 366

DB 455 MQPGSTFPRFPI-YSAALAKGM--TASTMINAPISLPKANGKAMNKNKNSDGRAGYI 511

QY 367 TLQYALQOSRNPVAVETLKNVGLNRKATFLNGIDIDPSIHYSNIAISSVTESDKKYGAS 426

DB 512 TLROALTASKKNWYSIRILMSIGIGYAOQYIQRGFKPSEIPASLMAIAGET-----T 565

QY 427 SEKMAAAYAPANGCTGYYPYTHKVFVPSDSEKE-----ESNVGTRAMKETTYAMMD 480

DB 566 PLRIAGYVFNANGYKVAHVIDKYDSQGRIRAOQMLVAGENAPQIIDPRNAVITMK 625

QY 481 MKTVELTYGTGRN-AYLAWLPQAGKTGTSNYPDEEIEHNHKTQSFPAPDELPAGYTRKYS 539

DB 626 INQDVVRVTAAGAAALGSDLAGKTGTIN-----DN-----KRAMYGVGFNPVY 670

QY 540 MAVMTGYSN-RLTPVNGCLTVAAKYRSMMTYLSGSPNEDMNIPEGLRNGEYFKNG 598

DB 671 TAVYIGFDPRSMGRAGYGTIAVPWVEYIGFALKGTGVKPMKAPEGVVTNGEYV--- 727

QY 599 AASTWSPAPQPPSPRESSSSDSTSSQSSSTPSTNNSTTTPNPNNTQGSNTTPOON 658

DB 728 -----MKRMTTSSDLADNSGIRPPTQPA--RRAVPNENRRRAESNTAPAREE 775

QY 659 QNPQAPQ 666

DB 776 SDFTPLP 783

RESULT 7

PPPA\_PSEAE STANDARD; PRT; 822 AA.

ID PPBA\_PSEAE

AC 007806;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Penicillin-binding protein 1A (PBP-1a) (PBP1a) [Includes: Penicillin-insensitive transglycosylase (EC 2.4.2.-) (peptidoglycan TGase)];

DE Penicillin-sensitive transpeptidase (EC 3.4.-.-) (Dd-transpeptidase)].

GN MRCA OR PONA OR PA5045.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI\_TaxId=287;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 15692 / PAOI;

Handfield J., Gagnon L., Dargis M., Huletsky A.;

OY	61	LQSNLSLQ--GGSTLTQOLIKLTLEFSTSTSDOTISKRAQEMALAIQLEQKATKOIILTYLIN	119
OY	109	LKSGHIDGGSTITIMOVAK--NYF--LTNESFSKRIEILIAQIEBRLTKDEILLEYNA	165
OY	120	KVYMSNGCYAGQCPAONVYCGKDLNNLSLPOLLACGMPAONQDYPSPHEAODERNLV	179
OY	166	KTYGNRRAYGIEAAQVYTKGPIKDLSLAEAMAGLKPASRYNPLVNFSTERRNNI	225
OY	180	LSEMKNOGYISAEQYERAVNPIITDGLQSLSKASNPAYMDNLKEVINOVEEET--GY	236
OY	226	LEHMKLGFIDQOQYCAAVEEPINASTHYQTPELMAYIAEMARALEWGVGSEATBEY	285
OY	237	NLLTGGMDVYTN-----VDQEAQ-----KHL-----	257
OY	286	KVITTVRSDDLQNASQSVRDGLIDYDQRHGYRGPETRLPGQTRDAMLKHLGQORSIGLE	345
OY	288	-----WD-----IYNTDEVAAIPDELOY	276
OY	346	PAIVTOYEKSGIMVYTRDQKEEAVYTMOSMKWAPFLSNNSGMPROPADVAQAQDILV	405
OY	277	-----ASTIVDVSNGKYLALQGA--RHQSSNVFGINQAVETNDDMS	317
OY	406	QROEDGTLRFVQIQAQSALISLDPKGCALRSLVGSGFSQSN-----YNRAIDAKRQPS	461
OY	318	TMKPIITDYAPALEXYGVYDSTATIVHDEPYN-----PGTNPVYNMDRGYFGN	365
OY	462	SEKPEI--YSALDNGF--TAASLVNDAPIVFVDEYLDKVMRPKNDENT-----FLGP	510
OY	366	ITLOALQOOSNNVAVETLNVGLNRAKKTFLNGIGIDYPSI--HYSNAISSNTTESDKKY	423
OY	511	IPLRALYKLSNNMYSIRPLQGLGIERAISYITKFGFQDELPRNFSIALGSTAT--	563
OY	424	GASSEKMAAAAFANGANGTYTKPKMYIKHVVFSDS-----SEKEFSN	464
OY	564	-VTPEIETIGASVYRANGYKKNPVIERISRQOQYLYQANPRVAYEEVDADADAQDN	622
OY	465	VG-----TRAKETAYMMTDMKIVLYTYGTR	492
OY	623	PGDEHPESAEGEGSIEAQVAAKQOTFEPTRPAERILIDARTAYIMTSMLODIKRGTR	682
OY	493	NAY--LAMLPRQKRGKTSYVDEIEENIKISQFAPARPELFAGTTRKYSAAVMTGYSNRLT	551
OY	663	RALALKRTDLAGKGTGN-----DSKDGWFSGYSNDSYVSWWGFPQPEP	727
OY	552	PLVGN--GLVVAKKVYRSMYTLSEGSNPEDMNIPGLGRNCEPFGKNGARSTWSPAP	608
OY	728	--LGRREGYGVALPIRIWYGFGLKKRPMHTMAEPGCI-----VSLRIDPTGSAAP	779
OY	609	QAPPESTESSSSSDSTSSQSSSTTPSTN	636
OY	780	GTGAYFEMFKND-----TPPSVN	799
RESULT	8		
PBPA_NEILA	STANDARD;	PRT;	798 AA.
ID	PBPA_NEILA		
AC	085759;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Penicillin-binding protein 1a (PBP-1a) (PBP1a) [includes: Penicillin-in-		
DE	insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan TGase)]		
DE	Penicillin-sensitive transpeptidase (EC 3.4.-.-) (bd-transpeptidase)]		
GN	MRA OR PONA.		
OS	Neisseria lactamica.		
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
OX	NCBI_TaxID=486;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NRL 3716;		
RA	Ropp P.A., Nicholas R.A.;		
RT	"Nucleotide sequence of the ponA gene encoding penicillin-binding		
RT	protein 1 from Neisseria lactamica.";		

```

RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC SUBUNITS) (BY SIMILARITY).
CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (by
CC similarity).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC TRANSGLYCOSYLASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRANSPEPTIDASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF085689; AAC35363.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl. 1.
KW Peptidoglycan synthesis; Cell wall; Transferrase; Glycosyltransferase;
KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
KW Signal-anchor; Antibiotic resistance.
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT DOMAIN 31 798 (POTENTIAL).
FT DOMAIN 50 218 PERIPLASMIC (POTENTIAL).
FT DOMAIN 378 700 TRANSGLYCOSYLASE.
FT ACI_SITE 461 461 ACYLATED BY PENICILLIN (BY SIMILARITY).
SQ SEQUENCE 798 AA; 88108 MW; 389C7672B86935D6 CRC64;

```

```

Query Match 18.48; Score 640; DB 1; Length 798;
Best Local Similarity 26.08; Pred. No. 1.6e-29;
Matches 203; Conservative 126; Mismatches 265; Indels 186; Gaps 24;

```

```

QY 2 IYDKKNOIALDGSERRNAQANDIPDLVKAISIEDHREDFRGIDITILILAFRLNL 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 52 VYSADGKIIGYGEORREFTKIGDPEVLRNAVIAEDKRFYQHMGVDVWGVARAVGNI 111
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 QSNELQ-GGSLTFOOLILTYFTSTSDQTSRKAQEAMLAIQLEOKATKOEILTYTINK 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 112 VSGSMQASASITITQOVAKNFYLS---SEKTFTRKFNELLLAKIKIDQSLSKILLEYLNK 168
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 VYKNGNGMOTAAQNYKQDLNLISLPQALLAGMPQAPNOQDPYSHPEAODRNLVL 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169 IYLGORAVGFAFAQIYFNKVDRLTLAEAMLAGLPKAPSAVNYVPERAKILQKYTL 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 SEMKNGGTSAEQYEKANPTITDGLQSLKSASNPAYMDN-----YLKEVING-VVEE 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 229 NNMLEEKITITQORDQALNEEL-----HYERFVQKIDQSALVVAEMVRELXEY 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 TGYNLITGTGMVDYTNV-----DOEAQ--- 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 278 YGEDAVYTGKLVYTVTRDHOKATEALRKALRNDRSSYRGAESYIDLDGEDAEAEAV 337
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 255 -KHLMDIYNDEYV-----AYPDE----- 273
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 338 SQYLSGLYTVDKMVPADVLTAKKKNVLIQPGCKRVTLDRRALCFARAVDNEMKEDR 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 274 -----LQVASTIYVSNQKVIADLGARHQSNNSPFGINQAVE 310
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 398 IIRGAVIRVRNNGRMVAVQEPPLQGLVSLDAKTGAVRALVGGYDFHSKT---FNRAIQ 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 311 TNRDWSMTKEITDYAPALEXYVDSTATIVHDEPYNPGIN-----TPVYNNDRGYFG 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 455 AMRQDGSFKEFV-YSAALSQGM--TASTMINDAPISLPKGPNGSVTP-KNSGRYSG 510
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 365 NITLQALQQRNRPVAVETLKKVGLNRAKTFNLNGIDIPSIHNSALSSNTTESDDKXG 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 511 YITLQALTLASKNNVSIILMSIGVYAQYIRFGRFPELPSLSALGTGE----- 565
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 425 ASSEKMAAAYAFANGCTGYRPMYHKVYFSDGSEKE-----PSNVGTRAMKETAYAM 478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 566 -TPRIAAIYFANGGRVSSYIDKTYDESGRLRAQMPQVLAQGNAPQAIIDPNNAIM 624
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 479 TDMKTVLTGTGRNA-YLAWLPQAGKTGNSYDEIEHNIKTSQVAPDELEFAGYTRK 537
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 625 YKIMQDVYRGVGTARGASALGRSDIAGKTGTIN-----DKAMFVGFNPD 669
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 538 YSMAMVTCYSN-RLTPLVNGLTVAAKYRSMYTLSSGSPEDMNIPGLYR-NGEEVF 595
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 670 VVTAVYIGFDRKPSMGKRGAGYGTIAVPWVDYMRPALKGRPGKMKMPDGVAVAGGEYIM 729
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 596 KN-----GARSTWNSPAPQGPS-----TESSSSSDSSTSSSTSTSTNNST 639
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 730 KEHWYTDPEGLMDNGGAAPQ--PSRKVVEDDGAAGGQGADESDRQDMQETVLPSTNT 787
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 9
PBPA_NEIMA STANDARD; PRT; 798 AA.
ID PBPA_NEIMA
AC 005194;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Penicillin-binding protein 1A (PBP-1a) (PBP1a) [includes: penicillin-
DE insensitive transglycosylase (EC 2.4.2.-) (peptidoglycan tgase);
DE penicillin-sensitive transpeptidase (EC 3.4.4.-) (Dd-transpeptidase)].
GN MRCA OR PONA OR NMA0653 OR NMB1807.
OS Neisseria meningitidis (serogroup A), and
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699, 491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z491 / Serogroup A / Serotype 4A;
RX MEDLINE=9725214; PubMed=9098083;
RA Ropp P.A., Nicholas R.A.;
RT "Cloning and characterization of the ponA gene encoding penicillin-
RT binding protein 1 from Neisseria gonorrhoeae and Neisseria
RT meningitidis.";
RL J. Bacteriol. 179:2783-2787(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tectelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.C., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Hart D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Clecko A., Parksey D.S., Blair E., Ciltone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathavan J.,

```





Query Match 18.2%; Score 633; DB 1; Length 798;  
 Best Local Similarity 26.3%; Pred. No. 4.1e-29;  
 Matches 191; Conservative 115; Mismatches 248; Indels 166; Gaps 19;

2 IYDNKNOLIADLGSERFRRNAQANDIPDIDVKAIVSIEDHRRFDHGRIDITRILGAFRLN 61  
 DB 52 IYSADGEVIGIGEORRRETKIGDPEVLRNVAIEDKRFQOHGVDWYARAVNGMT 111

62 QSNLSIQ-GGSLTLTQOLILKLYEFSTSDQTSRKAOEAWLAIOLBKATKOEILTYIYNK 120  
 DB 112 VAGVGSGASTITQOVAKNFYLS---SEKTFRKFNELALAVIKIQSLSKDIIELYFNQ 168

121 VYMSGNGTMOATAONYYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAAQDRRLNVL 180  
 DB 169 IYLGRAAGFASAAQIYFNKDVRELTLEAVMALGIPKAPSAVNPDIYVERAKLRQKYL 228

181 SEMKNGYISAEQYKAVNTPTDGLQSLKSASNPAYMDN-----YKKEYINQ-VEEE 233  
 DB 229 NNMLEKMTITLQQRQALNEEL-----HIERVQKIDQSALYAEVNRQELYRK 277

234 TGYNLLTGMVYTNVDOEAK----- 255

278 YGEDAVYTGQFKYTTVRDHDQKATEALRKALRNFDROSSYRGASYSIDLSKGEVEETV 337  
 DB 256 --HLMDIYNTDE----- 265

338 SOYLSGLTYTVDKMPAIVLDVTKRRNVYIOLPSGKRVTLIDGSLGFAARAANNKMGESR 397  
 DB 266 -----YVAVPDEELQVASTIVDSNGKYIAQAGARHOSNSVFGINQAE 310

388 IRRGSVIRNRNGGKMWVVOEPLLOATLVSLDAKTGAVALGVDGFHSKT---FNRAAQ 454  
 DB 311 TNRDWSGTMKPTTDYAPALEYGVYDSTATIYVDEPYNYP--GTNTPVY---NMDRGYEFGN 365

455 AMRGQSGFFKPEI-YSAALSKGM--TASTMVDAPISLPGKANKSVMPKPSDGRISY 511  
 DB 366 ITLOYALQOORNPVAVETLTKVGLNRAKTFELNGLIDIPSHYNAISSNTTESDKKYGA 425

512 IYLRQALASKNMVSIIRILMSIGVYAHGYIORFGFKSELPASLSMALGTGET----- 565  
 DB 426 SSEKMAAAYAFANGCTGYKKPMYIHKYVFSQSEKE-----FSNVGTPRAMKETTAYMMF 479

566 TPLKTAIEIYVFANGGYRVSSHVIDKTYGSDRLRAQOMQPLVAGONAPADIPRANAYMY 625  
 DB 480 DMKTKVLTLYGTGRN-AYLAMLPOACKGTGTSNTDEIENHIKTSOFVAPDELFAGYTRKY 538

626 KIMQDVVAVGTARGAALGRSDIAGKTGTN-----DN-----KDAWFGFNPV 670  
 DB 539 SKAVWTGYSN-RLPPLVNGLTVAAKVYRSMNTYLSGGSNPEDMNIPEGLY-RNGEYVFK 596

671 VTAAVYIGDPKPSKMGAGYGTIAVPVWYMRFAKLGQCGMKRYPGCVSSNGEYMYK 730  
 DB

RESULT 12  
 PBPA\_VIBCH STANDARD; PRT; 825 AA.  
 ID PBPA\_VIBCH  
 AC O9KNUS;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Penicillin-binding protein 1A (PBP-1a) (PBPA) [Includes: Penicillin-  
 insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan TGase);  
 DE Penicillin-sensitive transpeptidase (EC 3.4.-.-) (DD-transpeptidase)].  
 GN MRCA OR PONA OR VC2635.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 NC NCB1\_TaxID=666;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
 RX MEDLINE=20406633; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayem L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi L., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Mierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae.";  
 RL Nature 406:477-483(2000).  
 CC -I- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED  
 CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A  
 CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN  
 CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE  
 CC TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE  
 CC SUBUNITS) (BY SIMILARITY).  
 CC -I- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.  
 CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By  
 CC similarity).  
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE  
 CC TRANSGLYCOSYLASE FAMILY.  
 CC -I- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE  
 CC TRANSPEPTIDASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE004330; AAF95776.1; ALT\_INIT.  
 CC TIGR: VC2635;  
 DR InterPro: IPR001264; Transglycosyl.  
 DR Pfam: PF00912; Transglycosyl; 1.  
 DR Prodom: PD001895; Transglycosyl; 1.  
 KW Peptidoglycan synthetase; Cell wall; Transferase; Glycosyltransferase;  
 KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;  
 KW Signal-anchor; Antibiotic resistance; Complete proteome.  
 FT DOMAIN 1 6  
 FT TRANSMEM 7 27  
 FT SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)  
 FT (POTENTIAL)  
 FT PERIPLASMIC (POTENTIAL).  
 FT DOMAIN 28 825  
 FT TRANSGLYCOSYLASE.  
 FT TRANSPEPTIDASE.  
 FT ACT SITE 413 752  
 FT ACT SITE 471 471  
 FT ACYLATED BY PENICILLIN (BY SIMILARITY).  
 SQ SEQUENCE 825 AA; 91986 MW; 8A953DA2650B9D2 CRC64;

Query Match 17.7%; Score 617; DB 1; Length 825;  
 Best Local Similarity 24.3%; Pred. No. 3.6e-28;  
 Matches 200; Conservative 128; Mismatches 220; Indels 276; Gaps 26;

1 KIYDNKNOLIADLGSERFRRNAQANDIPDIDVKAIVSIEDHRRFDHGRIDITRILGAFRLN 60  
 DB 49 QVEFQDGLKLIAGFGERKRRIPKLEMPKELIAVATIEDSRVYEHGFPDITRAFAV 108

61 LQSNLSIQ-GGSLTLTQOLILKLYEFSTSDQTSRKAOEAWLAIOLBKATKOEILTYIYNK 119  
 DB 109 LNSGSASQASTITQOLANFNFLS---NEKKYWRKKEIFIIHIEQLSKOEILELYIN 165

120 KYMSGNGTMOATAONYYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAAQDRRLNVL 179  
 DB 166 KIYLGIRSVGVGAAGAAQAYRGKVKDLTGELIALINGLPAPSTMPDIYSVEBATRRNVV 225

180 LSEKNGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNYLKEV-----INOVEEET 234  
 DB 226 LQRMDEKTYITKAEYDAARAEPVLPKRYGAEIELNAP-----YVAEIAAMVVERGEBA 280

235 GYNLLTGMVYTNVDOE-----AOKHLMDI----- 260  
 DB 281 AY---TSGMNTYITVDKLRANQAALINLLAYDERHRYRGAERKELMGNVQNPANSSITQL 337





ID PBPB\_HAEMIN STANDARD: PRT: 853 AA.  
AC P31776;  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Penicillin-binding protein 1A (PBP-1a) (PBP1a) (Penicillin-binding protein A) (includes: Penicillin-insensitive transglycosylase DE (EC 2.4.2.-) (Peptidoglycan Gase); Penicillin-sensitive DE transpeptidase (EC 3.4.-.-) (DD-transpeptidase)).  
GN MCA OR POMA OR H10440.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OX Haemophilus.  
NCBI\_TaxID=727;  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN-RD / KW20 / ATCC 51907;  
RX MEDLINE=92009183; PubMed=1916268;  
RA Tomb J.-F., El-Hajj H., Smith H.O.;  
RT "Nucleotide sequence of a cluster of genes involved in the transformation of Haemophilus influenzae Rd.";  
RL Gene 104:1-10(1991).  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN-RD / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";  
RL Science 269:496-512(1995).  
RN  
RP CHARACTERIZATION.  
RC STRAIN-RD / KW20 / ATCC 51907;  
RX MEDLINE=96074300; PubMed=7592463;  
RA Sharma U.K., Dwarkanath P., Banerjee N., Town C., Balganesi T.S.;  
RT "Expression and characterization of the ponA (ORF I) gene of Haemophilus influenzae: functional complementation in a heterologous system.";  
RT Haemophilol. 177:6745-6750(1995).  
RN  
RP IDENTIFICATION BY MASS SPECTROMETRY.  
RX MEDLINE=20137488; PubMed=10675023;  
RA Laugen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B., Gray C., Fountoulakis M.;  
RT "Two-dimensional map of the proteome of Haemophilus influenzae.";  
RL Electrophoresis 21:411-429(2000).  
CC  
CC -1- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE SUBUNTS) (BY SIMILARITY).  
CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By similarity).  
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE TRANSGLYCOSYLASE FAMILY.  
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE TRANSPEPTIDASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL: M62809; AAA25007.1; -;  
DR EMBL: U32727; AAC22099.1; -;  
DR PIR: JH0438; JH0438.  
DR TIGR: H10440; -;  
DR InterPro: IPR001264; Transglycosyl.  
DR InterPro: IPR001460; Transpeptidase.  
DR Pfam: PF00912; Transglycosyl; 1.  
DR Pfam: PF00905; Transpeptidase; 1.  
DR ProDom: PD001895; Transglycosyl; 1.  
KW Peptidoglycan synthesis; Cell wall; 1. Transferrase; glycosyltransferase;  
KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;  
KW Signal-anchor; Antibiotic resistance; Complete proteome.  
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).  
FT TRANSSEM 7 27 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)  
FT  
FT DOMAIN 28 853 PERIPLASMIC (POTENTIAL).  
FT DOMAIN 37 205 TRANSGLYCOSYLASE.  
FT DOMAIN 387 681 TRANSPEPTIDASE.  
FT ACT\_SITE 441 441 ACYLATED BY PENICILLIN (BY SIMILARITY).  
FT CONFLICT 33 33 L -> LPSVERTLKTVE (IN REF. 1).  
SQ SEQUENCE 853 AA; 94221 MW; FC0846096CDB63B CRC64;  
  
Query Match 17.0%; Score 591.5; DB 1; Length 853;  
Best Local Similarity 29.2%; Pred. No. 1;le-26;  
Matches 166; Conservative 99; Mismatches 176; Indels 127; Gaps 20;  
  
QY 1 KIYDKKNIQDLIGSERVNAQNDIPDLYKATISIEDHREPHGIDTIRILGA-FLR 59  
Db 38 QITRDGKLTIGEGVGRKIPVKLADVPORLIDAFATEDSRFYDHGIDPGLIAALFYA 97  
QY 60 NLOSNLGGSTLTQOLIKLTYFTSTSDQTSRKAQAMLAIOLEKATNOEILTYVIN 119  
Db 98 VSNCGASQSGASTITQOLARNFPL--TSEKTIIRKAREAVLAVEIENTLNKOEILEYLN 154  
QY 120 KYVNSNGYGMQTAQNTYGGDLNLSLPOLALLAGMPQAPNOQPYSHPEAODRMLY 179  
Db 155 KIFLYRSYGVAAAOYFQKSLNLTLEMAIILGLPKASTMPLSLKSEERRNV 214  
QY 180 LSEKNGOYISAEOYERKVNPIPTDGLQSLKSASVPAVMQDYKEVINQVEEFGYML 239  
Db 215 LSRMLDEKTYISKEEDALKEPTVASYGAKFEPRADYVTEVROEMKRRGEENAY-- 271  
QY 240 TTGMQVYTNV--DQ-EAOK-----HLMD-----LYNTDEVYAV--- 269  
Db 272 TSGYVFETTVLSKQDAEQAKAVRNRLIDYDMHGYRGGAPLMQKNKAAAMDRIYGFIRK 331  
QY 270 -PDDELQYASTIVDSNG--KVIAQLGAR-----HOSN- 300  
Db 332 LPDSEPTIPAAVTIGVKGADILLASGKMKLSTNAMRWTKRSNPKVKGEOIWIHQANG 391  
QY 301 -----VSFG-----INQAVETNRDMSGTKPTITDAP 327  
Db 392 EMQLQIPANSALVSLNSDNGALFVAVGSEYDSKTRKATQSLVQVSSSTKPTI-FAA 450  
QY 328 ALEVGYDSTATIVHDEPNY--PG-----TNPFVYMDRQFGNITLQVALQDSRNP 379  
Db 451 ALERKL--TLSSVLQDPSITISQKPGQKMPKNSP---DR-YDSFPLRLRGLGSKNII 503  
QY 360 AVETLNKVLNRAKTFPLNGLGIDYPSIHYSNAISNTTESDKRYGASS---EKAAYAA 436  
Db 504 AIRAIGTIGIDFTAEFLRFGFRKRDQYFASLEIA-----LGAASFTPLEMARAYAV 554  
QY 437 FANGGYKKPMYIHKVVPDSSEKFEFN 464  
Db 555 FDNQGFLEPYITIEKIQNTGKOLFIA 582  
  
RESULT 15  
PBPB\_VIRCH



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2002, 04:24:19 ; Search time 86.16 Seconds

(without alignments)  
1337.217 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484  
Sequence: 1 KIYDNKNQIADLGSERRV.....TQSSNTTPDOQNQNPQAP 666

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19:\*

1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp rodent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriaph:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	719	2	Q9REU0 streptococc
2	3468	99.5	719	2	Q9RET9 streptococc
3	3403	97.7	719	2	Q9RET3 streptococc
4	3329	95.6	719	2	Q9WVW0 streptococc
5	3314	95.1	719	2	Q9RET8 streptococc
6	3299	94.7	719	2	Q9WVW1 streptococc
7	3274	94.0	719	2	Q9RET4 streptococc
8	3247	93.2	719	2	Q9RET6 streptococc
9	3246	93.2	719	2	Q9RET6 streptococc
10	3245	93.1	719	2	Q54947 streptococc
11	3243	93.1	719	2	Q54946 streptococc
12	3241	93.0	719	2	Q54948 streptococc
13	3231	92.7	719	2	Q9RET5 streptococc
14	3192	91.6	719	2	Q54950 streptococc
15	3134	90.0	718	2	Q9RET7 streptococc

17	2820	80.9	608	2	Q54951 streptococc
18	2791	80.1	728	2	Q9F2G7 streptococc
19	2696	77.4	727	2	Q9F2G4 streptococc
20	2432	69.8	527	2	Q70037 streptococc
21	2091	60.0	398	2	Q9R7M6 streptococc
22	2088	59.9	398	2	Q87102 streptococc
23	2037.5	58.5	721	16	Q99Y11 streptococc
24	2029	58.2	398	2	Q9R7M5 streptococc
25	2015	57.8	398	2	Q87108 streptococc
26	2011	57.7	398	2	Q87105 streptococc
27	1936	55.6	398	2	Q9R7M7 streptococc
28	1926	55.3	398	2	Q87103 streptococc
29	1921	55.1	398	2	Q9R7M8 streptococc
30	1911	54.9	793	2	Q9EXM9 enterococcu
31	1881	54.0	398	2	Q87107 streptococc
32	1876	53.8	398	2	Q87104 streptococc
33	1757	50.4	400	2	Q87106 streptococc
34	1751	50.3	778	2	Q9EXN1 enterococcu
35	1657.5	47.6	664	16	Q9C123 lactococcu
36	1592	45.7	310	2	Q52741 streptococc
37	1583	45.4	310	2	Q52743 streptococc
38	1573	45.1	310	2	Q52739 streptococc
39	1571	45.1	310	2	Q52737 streptococc
40	1567	45.0	310	2	Q52738 streptococc
41	1564	44.9	310	2	Q52740 streptococc
42	1558	44.7	310	2	Q52742 streptococc
43	1454	41.7	310	2	Q93N72 streptococc
44	1449	41.6	310	2	Q52744 streptococc
45	1201	34.5	826	16	Q92AB6 streptococc

## ALIGNMENTS

RESULT	ID	Q9REU0	PRELIMINARY:	PRT:	719 AA.
AC	Q9REU0	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	PENICILLIN-BINDING PROTEIN 1A.				
GN	PBP1A.				
OS	Streptococcus pneumoniae.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1313;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-SP 1261;				
RA	MEDLINE=21432820; PubMed=11549185;				
RX	Ferroni A., Berche P.;				
RT	"Alterations to penicillin-binding proteins 1A, 2B and 2X amongst penicillin-resistant clinical isolates of Streptococcus pneumoniae"				
RT	serotype 23F from the nasopharyngeal flora of children."				
RL	J. Med. Microbiol. 50:828-832(2001).				
DR	EMBL; AF210745; AAF17255.1;				
DR	InterPro; IPR001264; Transglycosyl.				
DR	InterPro; IPR001460; Transpeptidase.				
DR	Pfam; PF00912; Transglycosyl; 1.				
DR	Pfam; PF00905; Transpeptidase; 1.				
DR	ProDom; PD001895; Transglycosyl; 1.				
DR	SEQUENCE 719 AA; 79758 MW; 5BD397E83B4B3AA6 CRC64;				

QY	1	KIYDNKNQIADLGSERRVNAQNDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
DB	54	KIYDNKNQIADLGSERRVNAQNDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113

OY	61	LOSNILOGSSTLTQOLIKLTYESTSTSPQTSIKRAQEMWLAIOLEOKATKOELLTYINK	120
Db	114	LOSNILOGSSTLTQOLIKLTYESTSTSPQTSIKRAQEMWLAIOLEOKATKOELLTYINK	173
OY	121	VYMSNGTGMQTAAGNYGYGKDLNLSLPQALLAGMPAPNOQDYPYSHPEAADRRNLVL	180
OY	181	SEMKNQGYTSAQOYKAVANTPTTQSLQSLKASWYPAWMDYLUKEVINQVBEELGYMLLT	240
Db	234	SEMKNGYISAQOYKAVANTPTTQSLQSLKASWYPAWMDYLUKEVINQVBEELGYMLLT	293
OY	241	TGMDVYTNVDOQAQKHLMDIYNTDEYVYVPDDELQVASTIVDSNGKVIYAQLGARHOSN	300
Db	294	TGMDVYTNVDOQAQKHLMDIYNTDEYVYVPDDELQVASTIVDSNGKVIYAQLGARHOSN	353
OY	301	VSEFGINQAVETNRDMSGTMRKPTTDAPARLLEYGYDSTATIYHDEPRYVPGTNPVYWMDR	360
Db	354	VSEFGINQAVETNRDMSGTMRKPTTDAPARLLEYGYDSTATIYHDEPRYVPGTNPVYWMDR	413
OY	361	GVEFGNTLTQYALQOSRNPAYETLUKVLNRAKTFNLGIDGYSIHYSNAISSNTTESD	420
Db	414	GVEFGNTLTQYALQOSRNPAYETLUKVLNRAKTFNLGIDGYSIHYSNAISSNTTESD	473
OY	421	KRYGASSEKMAAAYAAPANGTYRPMYIHKVPESDSEKEFSNVGTRAKETTAAYMTD	480
Db	474	KRYGASSEKMAAAYAAPANGTYRPMYIHKVPESDSEKEFSNVGTRAKETTAAYMTD	533
OY	481	MKKTYLTGTGNNALAWLPQAGKGTGINSYTDDEIENHIKTSQCPVADDELFCATYTRYSM	540
Db	534	MKKTYLTGTGNNALAWLPQAGKGTGINSYTDDEIENHIKTSQCPVADDELFCATYTRYSM	593
OY	541	AVMTGYSNRLTPLVNGSLTVAAKYRSMNTYLSGSGNPEDMNIPEGLYRNGEEVFKNGAR	600
Db	594	AVMTGYSNRLTPLVNGSLTVAAKYRSMNTYLSGSGNPEDMNIPEGLYRNGEEVFKNGAR	653
OY	601	SWMNSPAPQOPSTSTSSSSSDSSTSSQSSSTTBSTNNSTTTPNNNTQQSNTTPDOONON	660
Db	654	SWMNSPAPQOPSTSTSTSSSSSDSSTSSQSSSTTBSTNNSTTTPNNNTQQSNTTPDOONON	713
OY	661	PQPAOP 666	
Db	714	PQPAOP 719	
RESULT	2		
Q9RET9			
ID	Q9RET9	PREDIMINARY;	PRT; 719 AA.
AC	Q9RET9;		
DT	01-MAY-2000 (Tremblrel. 13, Created)		
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)		
DT	01-DEC-2001 (Tremblrel. 13, Last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A.		
GN	PBP1A.		
OS	Streptococcus pneumoniae.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
CC	Streptococcus.		
OX	NCBI_TaxID=1313;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=SP 1513;		
RX	MEDLINE=21432820; PubMed=11549185;		
RA	Ferroni A., Berche P.;		
RT	"Alterations to penicillin-binding proteins 1A, 2B and 2X amongst		
RT	penicillin-resistant clinical isolates of Streptococcus pneumoniae		
RT	serotype 23F from the nasopharyngeal flora of children."		
RL	J. Med. Microbiol. 50:828-833(2001).		
DR	EMBL; AF210746; AAF17256.1; "		
DR	InterPro: IPR001264; Transglycosyl.		
DR	InterPro: IPR001460; Transpeptidase.		
DR	Pfam; PF00912; Transglycosyl. 1.		
DR	Pfam; PF00905; Transpeptidase. 1.		
DR	Pfam; PD001895; Transglycosyl.1.		

SEQ	SEQUENCE	719 AA:	79830 MW:	5f6776b8dbae3840 CRC64:
QY	1	KIYDNKNOLLADLGSRRRRNAQANDIPTDLVVAIVSIEDHREFDRHGIDITITILGAFLEFN	60	
QY	54	KIYDNKNOLLADLGSRRRRNAQANDIPTDLVVAIVSIEDHREFDRHGIDITITILGAFLEFN	113	
QY	61	LOSNSLGGSTLTLOOLIKLTYFSTSTSDOTSISKAQEAALALQLEKATKQEIILTYINK	120	
Db	114	LOSNSLGGSTLTLOOLIKLTYFSTSTSDOTSISKAQEAALALQLEKATKQEIILTYINK	173	
QY	121	VYMSKNYGMQRAAQNRYCKDLNNLSLPOLLALHAPQAPNPDYPSHPEAAQDRRNLYL	180	
Db	174	VYMSKNYGMQRAAQNRYCKDLNNLSLPOLLALHAPQAPNPDYPSHPEAAQDRRNLYL	233	
QY	181	SEKKNNGYISADQYEKAVNPTITDGLQSLSKASNPAYADNLTKEVINQVEEETGYNLT	240	
Db	234	SEKKNNGYISADQYEKAVNPTITDGLQSLSKASNPAYADNLTKEVINQVEEETGYNLT	293	
QY	241	TGMDVYTNVDOEAQKHLMDIYNTDEVAYAPDDELQVASTIVDSNGKYIAQLGARHSSN	300	
Db	294	TGMDVYTNVDOEAQKHLMDIYNTDEVAYAPDDELQVASTIVDSNGKYIAQLGARHSSN	353	
QY	301	VSPGIAQAVETNRDMSGTMRPTITDVAAPALEGVYSTATIVHDEPNYNGTTPVYNNDR	360	
Db	354	VSPGIAQAVETNRDMSGTMRPTITDVAAPALEGVYSTATIVHDEPNYNGTTPVYNNDR	413	
QY	361	GYPGNITLLOVALAQQSSRNPAVETLKNVGNRAKTFELNGIDYPSIHNASNAISSNTTESD	420	
Db	414	GYPGNITLLOVALAQQSSRNPAVETLKNVGNRAKTFELNGIDYPSIHNASNAISSNTTESD	473	
QY	421	KRYGASSEMAAAYAAAFANGGTYTPMYTHKRYVSDGSEKESNNGTRAMKETTAYMMTD	480	
Db	474	KRYGASSEMAAAYAAAFANGGTYTPMYTHKRYVSDGSEKESNNGTRAMKETTAYMMTE	533	
QY	481	MKTYVLTYGTGNATLAWLPQAGKGTGSNYTDEELENHKTQSOPAPDELFGYTRYKSM	540	
Db	534	MKTYVLTYGTGNATLAWLPQAGKGTGSNYTDEELENHKTQSOPAPDELFGYTRYKSM	593	
QY	541	AVMTGYSNRLPLVNGLTVAAKVYRSMTVYLSEGSNPEDMNIPGLYRNGEFVKNGAR	600	
Db	594	AVMTGYSNRLPLVNGLTVAAKVYRSMTVYLSEGSNPEDMNIPGLYRNGEFVKNGAR	653	
QY	601	SWNNSPAPQOPRSTSSSSSSSSSTSSSSSTPPTNNNTTTPPNNNTQOOSNTTPOOON	660	
Db	654	SWNNSPAPQOPRSTSSSSSSSSSTSSSSSTPPTNNNTTTPPNNNTQOOSNTTPOOON	713	
QY	661	PQPAQP 666		
Db	714	PQPAQP 719		
RESULT	3			
Q9RETS	ID	PRELIMINARY:	PRT:	719 AA.
AC	Q9RETS3:			
DT	01-MAY-2000 (Tremblrel. 13, Created)			
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)			
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)			
DE	PENICILLIN-BINDING PROTEIN 1A.			
GN	PBP1A.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Streptococcus			
OX	NCBI_TaxID=1313;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BM 4200;			
RX	MEDLINE=21432820; PubMed=11549185;			

RA Ferroni A., Berche P.;  
RT "Alterations to penicillin-binding proteins 1a, 2b and 2x amongst  
penicillin-resistant clinical isolates of *Streptococcus pneumoniae*  
serotype 23F from the nasopharyngeal flora of children.";  
J. Med. Microbiol. 50:828-832(2001).  
RL EMBL: AF210752; AAF17262.1; -.  
DR InterPro: IPR001264; Transglycosyl.  
DR InterPro: IPR001460; Transpeptidase.  
DR Pfam: PF00912; Transglycosyl. 1.  
DR Pfam: PF00905; Transpeptidase. 1.  
DR ProDom: PD001895; Transglycosyl. 1.  
SQ SEQUENCE 719 AA; 79781 MW; DB5993f6BD47F72D CRC64;

Query Match 97.7%; Score 3403; DB 2; Length 719;  
Best Local Similarity 97.3%; Pred. No. 5.4e-177;  
Matches 648; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 KIYDKKNOLLADLGSERVNQAANDIPDLYKAIVSIEDHREFDHRGIDTIRILGAFLRN 60  
DB 54 KIYDKKNOLLADLGSERVNQAANDIPDLYKAIVSIEDHREFDHRGIDTIRILGAFLRN 113  
QY 61 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAEAMLAIQLEOKATKOEILTYIYNK 120  
DB 114 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAEAMLAIQLEOKATKOEILTYIYNK 173  
QY 121 YVMSNGNGMGTAAONYYGKDLNNLSLPOLALLAGMPAPNOYDXYSHPEAAQDRNLVL 180  
DB 174 YVMSNGNGMGTAAONYYGKDLNNLSLPOLALLAGMPAPNOYDXYSHPEAAQDRNLVL 233  
QY 181 SEMKNOGYTISAQYKAVNTPITDGLQSLKSASNPAYMDNLYKVINOVEETGYNLT 240  
DB 234 SEMKNOGYTISAQYKAVNTPITDGLQSLKSASNPAYMDNLYKVINOVEETGYNLT 293  
QY 241 TGMDDYTVNDQEAQKHLMDIYNDEYVAYPDDELQVASTIVDVNSGKYIAQLGARHOSN 300  
DB 294 TGMDDYTVNDQEAQKHLMDIYNDEYVAYPDDELQVASTIVDVNSGKYIAQLGARHOSN 353  
QY 301 VSEGINQAVETNRDWSGTMKPTTDYAPALEGVYDSTATIYHDEBYNTPGINTPYVNMNR 360  
DB 354 VSEGINQAVETNRDWSGTMKPTTDYAPALEGVYDSTATIYHDEBYNTPGINTPYVNMNR 413  
QY 361 GYFGNITLQYALQOOSRNPAVETLNKVGILNRAKPTFLNGLGIDYPSIHNSNAISSWTESD 420  
DB 414 GYFGNITLQYALQOOSRNPAVETLNKVGILNRAKPTFLNGLGIDYPSIHNSNAISSWTESD 473  
QY 421 KKYGASSEKMAAAYAAAFANGGTYRKPMYIHKVYFSDGSEKESNVGTAMKETTAYMMTD 480  
DB 474 KKYGASSEKMAAAYAAAFANGGTYRKPMYIHKVYFSDGSEKESNVGTAMKETTAYMMTD 533  
QY 481 MKKTYLVLTGTRNAYLAWLPQAGKTGTSNYTDEELENHIKTSQVAPDELPAGYTRKYSM 540  
DB 534 MKKTYLVLTGTRNAYLAWLPQAGKTGTSNYTDEELENHIKTSQVAPDELPAGYTRKYSM 593  
QY 541 AVMTGYSNRLPLVGNGLTVAAKYVRSMMTYLSSEGNPEDMNIPEGILRNGEFVYKKNAR 600  
DB 594 AVMTGYSNRLPLVGNGLTVAAKYVRSMMTYLSSEGNPEDMNIPEGILRNGEFVYKKNAR 653  
QY 601 STWNSPAPQPPSTRESSSSSDSSSTSPSTNNSTTNPNNNTQOOSNTTPOQONON 660  
DB 654 STWNSPAPQPPSTRESSSSSDSSSTSPSTNNSTTNPNNNTQOOSNTTPOQONON 713  
QY 661 POPAP 666  
DB 714 POPAP 719

RESULT 4  
Q9WVW0 PRELIMINARY; PRT; 719 AA.  
AC Q9WVW0;  
DT 01-NOV-1999 (TRENBLREL. 12; Created)  
DT 01-NOV-1999 (TRENBLREL. 12; Last sequence update)

DT 01-OCT-2001 (TRENBLREL. 18; Last annotation update)  
DE PENICILLIN-BINDING PROTEIN 1A.  
GN BPPIA.  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
NM NCBI\_TaxID=1313;  
RN [1]

RC STRAIN=PO-342, PO-273, AND PO-341;  
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;  
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V  
clone of *Streptococcus pneumoniae* arose by large recombinational  
replacements of the cpsA-bppla region.";  
RT Microbiology 0:0-0(1999).

DR EMBL: AF139886; AAD43069.1; -.  
DR EMBL: AF139884; AAD43067.1; -.  
DR EMBL: AF139885; AAD43068.1; -.  
DR InterPro: IPR001264; Transglycosyl.  
DR InterPro: IPR001460; Transpeptidase.  
DR Pfam: PF00912; Transglycosyl. 1.  
DR Pfam: PF00905; Transpeptidase. 1.  
DR ProDom: PD001895; Transglycosyl. 1.  
SQ SEQUENCE 719 AA; 79622 MW; 7FBA6A75EA8FF8B3 CRC64;

Query Match 95.6%; Score 3329; DB 2; Length 719;  
Best Local Similarity 94.4%; Pred. No. 5.6e-173;  
Matches 629; Conservative 25; Mismatches 12; Indels 0; Gaps 0;

QY 1 KIYDKKNOLLADLGSERVNQAANDIPDLYKAIVSIEDHREFDHRGIDTIRILGAFLRN 60  
DB 54 KIYDKKNOLLADLGSERVNQAANDIPDLYKAIVSIEDHREFDHRGIDTIRILGAFLRN 113  
QY 61 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAEAMLAIQLEOKATKOEILTYIYNK 120  
DB 114 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAEAMLAIQLEOKATKOEILTYIYNK 173  
QY 121 YVMSNGNGMGTAAONYYGKDLNNLSLPOLALLAGMPAPNOYDXYSHPEAAQDRNLVL 180  
DB 174 YVMSNGNGMGTAAONYYGKDLNNLSLPOLALLAGMPAPNOYDXYSHPEAAQDRNLVL 233  
QY 181 SEMKNOGYTISAQYKAVNTPITDGLQSLKSASNPAYMDNLYKVINOVEETGYNLT 240  
DB 234 SEMKNOGYTISAQYKAVNTPITDGLQSLKSASNPAYMDNLYKVINOVEETGYNLT 293  
QY 241 TGMDDYTVNDQEAQKHLMDIYNDEYVAYPDDELQVASTIVDVNSGKYIAQLGARHOSN 300  
DB 294 TGMDDYTVNDQEAQKHLMDIYNDEYVAYPDDELQVASTIVDVNSGKYIAQLGARHOSN 353  
QY 301 VSEGINQAVETNRDWSGTMKPTTDYAPALEGVYDSTATIYHDEBYNTPGINTPYVNMNR 360  
DB 354 VSEGINQAVETNRDWSGTMKPTTDYAPALEGVYDSTATIYHDEBYNTPGINTPYVNMNR 413  
QY 361 GYFGNITLQYALQOOSRNPAVETLNKVGILNRAKPTFLNGLGIDYPSIHNSNAISSWTESD 420  
DB 414 AFGNITLQYALQOOSRNPAVETLNKVGILNRAKPTFLNGLGIDYPSIHNSNAISSWTESD 473  
QY 421 KKYGASSEKMAAAYAAAFANGGTYRKPMYIHKVYFSDGSEKESNVGTAMKETTAYMMTD 480  
DB 474 KKYGASSEKMAAAYAAAFANGGTYRKPMYIHKVYFSDGSEKESNVGTAMKETTAYMMTD 533  
QY 481 MKKTYLVLTGTRNAYLAWLPQAGKTGTSNYTDEELENHIKTSQVAPDELPAGYTRKYSM 540  
DB 534 MKKTYLVLTGTRNAYLAWLPQAGKTGTSNYTDEELENHIKTSQVAPDELPAGYTRKYSM 593  
QY 541 AVMTGYSNRLPLVGNGLTVAAKYVRSMMTYLSSEGNPEDMNIPEGILRNGEFVYKKNAR 600  
DB 594 AVMTGYSNRLPLVGNGLTVAAKYVRSMMTYLSSEGNPEDMNIPEGILRNGEFVYKKNAR 653  
QY 601 STWNSPAPQPPSTRESSSSSDSSSTSPSTNNSTTNPNNNTQOOSNTTPOQONON 660  
DB 654 STWNSPAPQPPSTRESSSSSDSSSTSPSTNNSTTNPNNNTQOOSNTTPOQONON 713

QY 661 POPAP 666  
Db 714 POPAP 719

RESULT 5  
Q9RET8 PRELIMINARY; PRT; 719 AA.  
AC Q9RET8; TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
DE PENICILLIN-BINDING PROTEIN 1A.  
GN PBPIA.  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae.  
OC Streptococcus.  
OX NCBI\_taxid=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SP 1465;  
RX MEDLINE=21432820; PubMed=11549185;  
RA Ferroni A., Berche P.;  
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst penicillin-resistant clinical isolates of Streptococcus pneumoniae serotype 23F from the nasopharyngeal flora of children.";  
RL J. Med. Microbiol. 50:828-832(2001).  
DR EMBL: AF210747; AAF17257.1; -  
DR InterPro: IPR001264; Transglycosyl.  
DR InterPro: IPR001460; Transpeptidase.  
DR Pfam: PF00905; Transglycosyl; 1.  
DR ProDom: PD001895; Transglycosyl; 1.  
SQ SEQUENCE 719 AA; 79662 MW; A7B96EA15434A805 CRC64;

Query Match 95.1%; Score 3314; DB 2; Length 719;  
Best Local Similarity 94.1%; Pred. No. 3,7e-172;  
Matches 627; Conservative 27; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTYDNKQNLADIGSERVNAQANDIPTDLVKAIVSTEDHREFDHRGIDTRITLGAFLRN 60  
Db 54 KTYDNKQNLADIGSERVNAQANDIPTDLVKAIVSTEDHREFDHRGIDTRITLGAFLRN 113  
QY 61 LOSNSLOGGSLTLOQLIKLTFSTSTSDQTSRKAQEWLAIOLEOKATQOELITYINK 120  
Db 114 LOSNSLOGGSLTLOQLIKLTFSTSTSDQTSRKAQEWLAIOLEOKATQOELITYINK 173  
QY 121 VYNSNGNYGMQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVL 180  
Db 174 VYNSNGNYGMQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVL 233  
QY 181 SEMNNGYISAEQEKAVNPITDGLSLSASNPAYMNYLKEVINOVEEELGYMLT 240  
Db 234 SEMGQGYISAEQEKAVNPITDGLSLSASNPAYMNYLKEVINOVEEELGYMLT 293  
QY 241 TGMVYTNVDOEAKHMDIYNTEDEYVAYPPDELQVASTIYDVNSGVYIAQLGARHOSN 300  
Db 294 TGMVYTNVDOEAKHMDIYNTEDEYVAYPPDELQVASTIYDVNSGVYIAQLGARHOSN 353  
QY 301 VSEGINQAVETNRDWSGTMRPTIDYAPALEGYVDSTATIVHDEPPYNGPTNPEVYMWDR 360  
Db 354 VSEGINQAVETNRDWSGTMRPTIDYAPALEGYVDSTATIVHDEPPYNGPTNPEVYMWDR 413  
QY 361 GYRGNITLQALQOSRVNPAVETLNKGLNRAKTFGLNGIGDIPSHYSNAISSNTTESD 420  
Db 414 AYFGNITLQALQOSRVNPAVETLNKGLNRAKTFGLNGIGDIPSHYSNAISSNTTESN 473  
QY 421 KRYGASSEKMAAFAAFANGSTGYKPMYIKRVVPSDSESEKFEFNVGRANKETAYAMWT 480  
Db 474 KRYGASSEKMAAFAAFANGSTGYKPMYIKRVVPSDSESEKFEFNVGRANKETAYAMWT 533

QY 481 MKKTIVLGYTGRNAYLAWLPQAGKTGTSNXYDEEIEHNKTSQEVAPDELFAGYTRYSM 540  
Db 534 MKKTIVLGYTGRNAYLAWLPQAGKTGTSNXYDEEIEHNKTSQEVAPDELFAGYTRYSM 593  
QY 541 AVWTGYSNRLTPLYGNGITVAAKYRSMYTLSSGSNPEDMNTEGLYRNEEFKKGAR 600  
Db 594 AVWTGYSNRLTPLYGNGITVAAKYRSMYTLSSGSNPEDMNTEGLYRNEEFKKGAR 653  
QY 601 STWSSPAPQPPSTESSSSSDSSSTSSSTPSTNNSTTNNPNNNQOASWTFPDQONON 660  
Db 654 STWSSPAPQPPSTESSSSSDSSSTSSSTPSTNNSTTNNPNNNQOASWTFPDQONON 713  
QY 661 POPAP 666  
Db 714 POPAP 719

RESULT 6  
Q9WM11 PRELIMINARY; PRT; 719 AA.  
AC Q9WM11; TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)  
DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
DE PENICILLIN-BINDING PROTEIN 1A.  
GN PBPIA.  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae.  
OC Streptococcus.  
OX NCBI\_taxid=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-URU-E159; M134, URU-E135, AND URU-E157;  
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;  
RT "serotype 14 variants of the Spanish penicillin-resistant serotype 9V RT clone of Streptococcus pneumoniae arose by large recombinational RT replacements of the cpsA-pbpA region.";  
RL Microbiology 0:0-0(1999).  
DR EMBL: AF139887; AAD43073.1; -  
DR EMBL: AF139887; AAD43070.1; -  
DR EMBL: AF139888; AAD43071.1; -  
DR EMBL: AF139889; AAD43072.1; -  
DR EMBL: AF159448; AAD48430.1; -  
DR InterPro: IPR001264; Transglycosyl.  
DR InterPro: IPR001460; Transpeptidase.  
DR Pfam: PF00912; Transglycosyl; 1.  
DR Pfam: PF00905; Transpeptidase; 1.  
DR ProDom: PD001895; Transglycosyl; 1.  
SQ SEQUENCE 719 AA; 79701 MW; 3EDC8A21C94DAC87 CRC64;

Query Match 94.7%; Score 3299; DB 2; Length 719;  
Best Local Similarity 93.4%; Pred. No. 2,4e-171;  
Matches 622; Conservative 30; Mismatches 14; Indels 0; Gaps 0;

QY 1 KTYDNKQNLADIGSERVNAQANDIPTDLVKAIVSTEDHREFDHRGIDTRITLGAFLRN 60  
Db 54 KTYDNKQNLADIGSERVNAQANDIPTDLVKAIVSTEDHREFDHRGIDTRITLGAFLRN 113  
QY 61 LOSNSLOGGSLTLOQLIKLTFSTSTSDQTSRKAQEWLAIOLEOKATQOELITYINK 120  
Db 114 LOSNSLOGGSLTLOQLIKLTFSTSTSDQTSRKAQEWLAIOLEOKATQOELITYINK 173  
QY 121 VYNSNGNYGMQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHPEAODRRNLVL 180

Db	174	YVMSNGNQMGTAAQNYNYGKDLNLSLPQLALLAGMPAQPNQVDPSYSHPAQDQRRRLVLT	233
Qy	181	SEMNQGYISAEQYERKAVNPIITGLOSLKSASNYPAVMQNTLKEVYNQYEEETGYNLLT	240
Db	234	SEMNQGYISAEQYERKAVNPIITGLOSLKSASNYPAVMQNTLKEVYNQYEEETGYNLLT	233
Qy	241	TGMDYTTNVDEQAKHLMIDYINTEDEYVAAVPDDELQVASTIVDVSNGKVIAQLGARHOSSN	300
Db	294	TGMDYTTNVDEQAKHLMIDYINSEQYSYDPDDLQVASTIVDVSNGKVIAQLGARHOASN	353
Qy	301	VSEFGINQAVETNRDMSGTMRPIIDYAPALEXYVDSTATTIVHDEPNYPGNTPTVYNNMR	360
Db	354	VSEFGINQAVETNRDMSGTMRPIIDYAPALEXYVDSTATTIVHDEPNYPGNTPTVYNNMR	413
Qy	361	GYFGNITLQVLAQOOSRNPVAFETLNKXGLNRAKFEFLNGIGIDVPSIHYSAISNTTESD	420
Db	414	AIFGNTITLQVLAQOOSRNPVAFETLNKXGLNRAKFEFLNGIGIDVPSIHYSAISNTTESN	473
Qy	421	KRYGASSEKMAAAVAAAFANGTGYTKPKMYIHKVYFSDSSEKEFNSVNGVRANKETTAYVMTD	480
Db	474	KQYGASSEKMAAAVAAAFANGTGYTKPKMYIHKVYFSDGSKKEFSVNGVRANKETTAYVMTD	533
Qy	481	MMKTYLVYGTGTRNMYLMLPQAGKTGTSNTDEIENHITSOQVADDELFAGYTRKYSM	540
Db	534	MMKTYLVYGTGTRGAYLPLPQAGKTGTSNTDEVEHNIKNTGVADDEMFVGYTRKYSM	593
Qy	541	AVMTGYSNRLLPLVLGNGLTVAAKVYRRSMNTYLSEGSNPEDMNIPCEGLRNGEPEVFKNGAR	600
Db	594	AVMTGYSNRLLPLPIYGDGELVAAKVYRRSMITYLSDTHPEDMTMDGLFRNGEPEVFKNGAR	653
Qy	601	STWNSPAPQPPSTRESSSSSSSDSTSCOSSSTTPSTNNSTTTPNNNTQOOSNTTPDOONN	660
Db	654	STWNSPAPQPPSTRESSSSSSSDSTSCOSSSTTPSTNNSTTTPNNNTQOOSNTTPDOONN	713
Qy	661	POPAOP 666	
Db	714	POPAOP 719	
RESULT	7		
Q9RETA	Q9RETA	PRELIMINARY;	PRT: 719 AA.
AC	Q9RETA:		
DT	01-MAY-2000 (TReMBLrel. 13, Created)		
DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)		
DT	01-DEC-2001 (TReMBLrel. 19, Last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A.		
GN	PBP1A.		
OS	Streptococcus pneumoniae.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
CC	Streptococcus.		
OX	NCBI_TaxID=1313;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-SP 22861;		
RX	MEDLINE=21432820; PubMed=11549185;		
RA	Ferroni A., Berche P.;		
RT	"Alterations to penicillin-binding proteins 1a, 2b and 2x amongst		
RT	penicillin-resistant clinical isolates of Streptococcus pneumoniae		
RT	serotype 23F from the nasopharyngeal flora of children."		
RL	J. Med. Microbiol. 50:828-833(2001).		
DR	EMBL; AF210751; AAF17261.1; -		
DR	InterPro; IPR001264; Transglycosyl.		
DR	InterPro; IPR001460; Transpeptidase.		
DR	Pfam; PF009912; Transglycosyl. 1.		
DR	Pfam; PF009905; Transpeptidase. 1.		
DR	ProDom; PD001895; Transglycosyl. 1.		
QO	SEQUENCE 719 AA; 79567 MW; 7B6ECCBB8BCF286 CRC64;		

Query Match	94.0%;	Score 3274;	DB 2;	Length 719;
Best Local Similarity	92.9%;	Pred. No. 5.5e-170;		
Matches 619;	Conservative 31;	Mismatches 16;	Indels 0;	Gaps 0

QY	1	KIYNNKQNLADLDSERVVNAQANDIPDVLKAVISJEDHFPFHRIDTRIIGALRN	60
Db	54	KIYNNKQNLADLDSERVVNAQANDIPDVLKAVISJEDHFPFHRIDTRIIGALRN	1133
QY	61	LQSNLSGGSTLTQOLIKLTVFSTSTSDQTSIRKAEAMALALQLEOKATQOEILTVYINK	120
Db	114	LQSNLSGGSTLTQOLIKLTVFSTSTSDQTSIRKAEAMALALQLEOKATQOEILTVYINK	1737
QY	121	VYMSNGNVTGMOATAAQNYYGKDLNLSPQLALLAGMGAQAPNQPYPYSHPREAODRNLVL	180
Db	174	VYMSNGNVTGMOATAAQNYYGKDLNLSPQLALLAGMGAQAPNQPYPYSHPREAODRNLVL	2333
QY	181	SEMNKOGTISAEOYEKAVNPITPTGLOSLKSASNYPRYMDNYLKEVINOVDEEETGYMLT	240
Db	234	SEMNKOGTISAEOYEKAVNPITPTGLOSLKSASNYPRYMDNYLKEVINOVDEEETGYMLT	2933
QY	241	TGMDVYTVNDQEAOKHLMDIYNTDEYVAYPPDELEQVASTIVDVSNGKVLIAQLGARHOSN	300
Db	294	TGMDVYTVNDQEAOKHLMDIYNSDQYVSPDDDLQVASTVVDVSGKVLIAQLGARHOSN	353
QY	301	VSEGTINQAVETNRBWSGTMRKITDYADALEGYVDSTATIVYHDEPYNPQGTNPVYWMDR	366
Db	354	VSEGTINQAVETNRBWSGAMKPIITYDAPALIEGYVDSTATVYNDLPYNPQGTNPVYWMDR	4133
QY	361	GYFNGITLOVYALQOOSRNPVAVETLNKGLNKAFTFLNGLDIDPYSIYNSAISNTTESD	420
Db	414	AYFGNITLOVYALQOOSRNPVAVETLNKGLDRAKFTFLNGLDIDPYSIYNSAISNTTESN	4737
QY	421	KRYGASSEKMAAAYAAAFANGSTYYKPMYIHKVYFSDSESEKEFSNVGTRANKETTAYMMD	480
Db	474	KQYGSSEKMAAAYAAAFANGSIYHKPMYINKVYFSDSGKKEFSNVGTRANKETTAYMME	5333
QY	481	MMKVTVLYGTGRNAYLYAMLPAQKGTGNSNTDEIEHMITSQVVADELPAQTRKYSM	540
Db	534	MMKVTVLYAGTGRGAYLLPMLAQAGKGTGNSNTDDEIEHKIKNTGVADDEMVEGYTRKYSM	593
QY	541	AVMTGYSNRLTPELVNGCLTVAAKAYRSMTYTLSEGSNPEDMNIPEGLYRNGEYFKNGAR	600
Db	594	AVMTGYSNRLTPELVNGCLTVAAKAYRSMTYTLSEGSNPEDMNIPEGLYRNGEYFKNGAR	655
QY	601	STWNSPAPQOPPESTSSSSSDSSTSOSSSTTPSTNNSTTTPNPNNTQQSNTTTPDOONON	666
Db	654	STWSSPAPQOPPESTSSSSSDSSTSOSSNTTTPSTNNSTTTPNPNNTQQSNTTTPDOONON	7133
QY	661	PQPAQP 666	
Db	714	PQPAQP 719	
RESULT	8		
057114			
057114	PRELIMINARY:	PRT:	719 AA.
AC	057114:		
DT	01-NOV-1996 (Tremblrel. 01, Created)		
DT	01-NOV-1996 (Tremblrel. 01, Last sequence update)		
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A.		
GN	PEP1A.		
OS	Streptococcus pneumoniae.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=1313;		
OX	11)		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CS111 (CEPHALOSPORIN-RESISTANT CLINICAL ISOLATE);		
RA	MEDLINE=96012191; Pubmed=7574521;		
RA	Coffey T.J., Daniels M., McDougal L.K., Dowson C.G., Tenover F.C.,		
RA	Spencat B.G.;		
RT	"Genetic analysis of clinical isolates of Streptococcus pneumoniae		
RT	with high-level resistance to expanded-spectrum cephalosporins.";		
RL	Antimicrob. Agents Chemother. 39:1306-1313(1995).		
RL	EMBL: Z49095; CAA8918.1; -.		

DR EMBL; 249094; CAA88917.1; -  
DR InterPro: IPR001264; Transglycosyl.  
DR InterPro: IPR001460; Transpeptidase.  
DR Pfam: PF00912; Transglycosyl; 1.  
DR Pfam: PF00905; Transpeptidase; 1.  
DR Prodom: PD001895; Transglycosyl; 1.  
DR SEQUENCE 719 AA; 79712 MW; 5396607C0E67D06E CRC64;

Query Match 93.2%; Score 3247; DB 2; Length 719;  
Best Local Similarity 91.9%; Pred. No. 1.6e-166;  
Matches 612; Conservative 33; Mismatches 21; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADIGSERVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFLRN 60  
DB 54 KIYDNKNOLIADIGSERVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFLRN 113  
QY 61 LOSNLSGGSTLTQOLIKLTYFSTSDQTSIRKAOEAMLAIOEQKATKOEILTYINK 120  
DB 114 LOSNLSGGSTLTQOLIKLTYFSTSDQTSIRKAOEAMLAIOEQKATKOEILTYINK 173  
QY 121 VYMSNGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDYSHPPEAQRNRNVL 180  
DB 174 VYMSNGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDYSHPPEAQRNRNVL 233  
QY 181 SEMKNGYISAQYKAVNTPTDGLQSLKSASNPAYMDNLYKKEVINOVEETGYNLLT 240  
DB 234 SEMKNGYISAQYKAVNTPTDGLQSLKSASNPAYMDNLYKKEVINOVEETGYNLLT 293  
QY 241 TGMDDYTVNDQEAOKHLMIDYNTDEVAYPDDELQVASTYDVDSNGKYIAOLGARHQSNN 300  
DB 294 TGMDDYTVNDQEAOKHLMIDYNSDQYVSPDDDLQVASTYDVDSNGKYIAOLGARHQSNN 353  
QY 301 VSFGINQAVETNRDGMGSMKPTTDYAPALEYGVYSTATIVDEBYNPGTNPYYNMDR 360  
DB 334 VSFGINQAVETNRDGMGSMKPTTDYAPALEYGVYSTATIVDEBYNPGTNPYYNMDR 413  
QY 361 GYFGNITLQYALQOQRNRYPAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNMISSTTESD 420  
DB 414 VYFGNITLQYALQOQRNRYPAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNMISSTTESN 473  
QY 421 KKYGASSEKMAAAYAAPANGGTYKPKYIHKYVESDGSSEKESNVGTRAMKETTAYMMTD 480  
DB 474 KKYGASSEKMAAAYAAPANGGTYKPKYIHKYVESDGSSEKESNVGTRAMKETTAYMMTE 533  
QY 481 MKKTVLTGTCGNALVALPQAGKTGTSNYTDEETENHIKTSQFAPELFLAGYRRTKSM 540  
DB 534 MKKTVLTGTCGNALVALPQAGKTGTSNYTDEETENHIKTSQFAPELFLAGYRRTKSM 593  
QY 541 AVMTGYSNRLTPLVNGGLTVAAKYVRSMNTYLSSEGSNPEDNNIPGLYRNGEFYKNGAR 600  
DB 594 AVMTGYSNRLTPLVNGGLTVAAKYVRSMNTYLSSEDDQPODMWMPGLYRNGEFYKNGAR 653  
QY 601 STWNSPAPQPPSTSESSSSSDSSTSPSTNNSTTNNPNNNTQOQNTTPOQONON 660  
DB 654 STWNSPAPQPPSTSESSSSSDSSTSPSTNNSTTNNPNNNTQOQNTTPOQONON 713  
QY 661 POPAP 666  
DB 714 POPAP 719

RESULT 9

Q9RET6 PRELIMINARY; PRT; 719 AA.

AC Q9RET6; 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE PENICILLIN-BINDING PROTEIN 1A.  
GN PEB1A.  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SP 1053;  
RX MEDLINE=21432820; PubMed=11549185;  
RA Ferroni A., Berche P.;  
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst  
penicillin-resistant clinical isolates of Streptococcus pneumoniae  
serotype 23F from the nasopharyngeal flora of children.";  
RL J. Med. Microbiol. 50:828-832(2001).  
DR EMBL; AF210749; AAF17259.1; -  
DR InterPro: IPR001264; Transglycosyl.  
DR InterPro: IPR001460; Transpeptidase.  
DR Pfam: PF00912; Transglycosyl; 1.  
DR Pfam: PF00905; Transpeptidase; 1.  
DR Prodom: PD001895; Transglycosyl; 1.  
DR SEQUENCE 719 AA; 79638 MW; 1DAE3C5937048E16 CRC64;

Query Match 93.2%; Score 3246; DB 2; Length 719;  
Best Local Similarity 92.0%; Pred. No. 1.8e-166;  
Matches 613; Conservative 35; Mismatches 18; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADIGSERVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFLRN 60  
DB 54 KIYDNKNOLIADIGSERVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFLRN 113  
QY 61 LOSNLSGGSTLTQOLIKLTYFSTSDQTSIRKAOEAMLAIOEQKATKOEILTYINK 120  
DB 114 LOSNLSGGSTLTQOLIKLTYFSTSDQTSIRKAOEAMLAIOEQKATKOEILTYINK 173  
QY 121 VYMSNGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDYSHPPEAQRNRNVL 180  
DB 174 VYMSNGNGMGTAAQNYGKDLNNLSLPOLALLAGMPQAPNOYDYSHPPEAQRNRNVL 233  
QY 181 SEMKNGYISAQYKAVNTPTDGLQSLKSASNPAYMDNLYKKEVINOVEETGYNLLT 240  
DB 234 SEMKNGYISAQYKAVNTPTDGLQSLKSASNPAYMDNLYKKEVINOVEETGYNLLT 293  
QY 241 TGMDDYTVNDQEAOKHLMIDYNTDEVAYPDDELQVASTYDVDSNGKYIAOLGARHQSNN 300  
DB 294 TGMDDYTVNDQEAOKHLMIDYNSDQYVSPDDDLQVASTYDVDSNGKYIAOLGARHQSNN 353  
QY 301 VSFGINQAVETNRDGMGSMKPTTDYAPALEYGVYSTATIVDEBYNPGTNPYYNMDR 360  
DB 334 VSFGINQAVETNRDGMGSMKPTTDYAPALEYGVYSTATIVDEBYNPGTNPYYNMDR 413  
QY 361 GYFGNITLQYALQOQRNRYPAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNMISSTTESD 420  
DB 414 AVFGNITLQYALQOQRNRYPAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSNMISSTTESN 473  
QY 421 KKYGASSEKMAAAYAAPANGGTYKPKYIHKYVESDGSSEKESNVGTRAMKETTAYMMTD 480  
DB 474 KKYGASSEKMAAAYAAPANGGTYKPKYIHKYVESDGSSEKESNVGTRAMKETTAYMMTE 533  
QY 481 MKKTVLTGTCGNALVALPQAGKTGTSNYTDEETENHIKTSQFAPELFLAGYRRTKSM 540  
DB 534 MKKTVLTGTCGNALVALPQAGKTGTSNYTDEETENHIKTSQFAPELFLAGYRRTKSM 593  
QY 541 AVMTGYSNRLTPLVNGGLTVAAKYVRSMNTYLSSEGSNPEDNNIPGLYRNGEFYKNGAR 600  
DB 594 AVMTGYSNRLTPLVNGGLTVAAKYVRSMNTYLSSEDTHPEDTWPGLYRNGEFYKNGAR 653  
QY 601 STWNSPAPQPPSTSESSSSSDSSTSPSTNNSTTNNPNNNTQOQNTTPOQONON 660  
DB 654 STWNSPAPQPPSTSESSSSSDSSTSPSTNNSTTNNPNNNTQOQNTTPOQONON 713  
QY 661 POPAP 666  
DB 714 POPAP 719





```
|||||
Db 414 AFGNITTLQYALQOQRNRYAVETLNKVGIDRAKTEFLNGIDIDPSMHVANAISNTTESN 473
OY 421 KKYGASSEKMAAAYAFANGGTYTYPMTYHKVYFSDGSEKESNNGTRAMKETTAYMMTD 480
Db 474 KQYGASSEKMAAAYAFANGGTYTYPMTYHKVYFSDGSEKESNNGTRAMKETTAYMMTE 533
OY 481 MKKTVLTGTGNAYLAMLPOAGKTGTSNYTDEEIEHNHKTISOFAVAPDELFAGYTRKYSM 540
Db 534 MKKTVLTGTGNAYLAMLPOAGKTGTSNYTDEEIEHNHKTISOFAVAPDELFAGYTRKYSM 593
OY 541 AVWTGYSNRLPPLVGNGLTVAAKYRSMYTYLSEGSNPEDMNIPGLYRNGEVEFKNGAR 600
Db 594 AVWTGYSNRLPPLVGNGLTVAAKYRSMYTYLSEGSNPEDMNIPGLYRNGEVEFKNGAR 653
OY 601 STWNSPAPQOQPESTSSSSSSSTSSSTSPSTNNSTTTPNNNTQOOSNTTPOQONON 660
Db 654 STWNSPAPQOQPESTSSSSSSSTSSSTSPSTNNSTTTPNNNTQOOSNTTPOQONON 713
OY 661 POPAP 666
Db 714 POPAP 719

RESULT 12
O54946 PRELIMINARY; PRT: 719 AA.
AC 054946:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA OR PPPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OC NCBI_taxid=1313;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-456;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain."
RL EMBO J. 11:3831-3836(1992).
RN 12
RP SEQUENCE OF 310-619 FROM N.A.
RC STRAIN-8303, AND 35193;
RA Smith A.M., Klugman K.P.;
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
RN 13
RP SEQUENCE OF 320-717 FROM N.A.
RC STRAIN-#17/246, #20/898, AND #27/SHA3;
RX MEDLINE=98409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a thr-371 substitution in a conserved amino acid motif
RT of penicillin-binding protein 1A with penicillin resistance of
RT Streptococcus pneumoniae."
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
RN 14
RP SEQUENCE FROM N.A.
RC STRAIN-SP-665;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
RT clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-ppIIa region."
RL Microbiology 0:0-0(1999).
DR EMBL: X67868; CAA48068.1; -
DR EMBL: AF046230; AAC24695.1; -
DR EMBL: AB006876; BAA32072.1; -
DR EMBL: AF139883; AAD3066.1; -
DR InterPro: IPR001264; Transglycosyl.
```

```
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 719 719
SQ SEQUENCE 719 AA: 79736 MW: F090F91522C834DC CRC64:

Query Match 93.1%; Score 3243; DB 2; Length 719;
Best Local Similarity 91.9%; Pred. No. 2,7e-168;
Matches 612; Conservative 32; Mismatches 22; Indels 0; Gaps 0;

OY 1 KIYDKNQLIADLSEBRVNAQANDIPDLVKAIVSIEDHREDFRGIDITRILGAFLEN 60
Db 54 KIYDKNQLIADLSEBRVNAQANDIPDLVKAIVSIEDHREDFRGIDISRIILGAFLEN 113
OY 61 LOSNLSGGSTLTQOLIKLTYFSTSTSDQTSIRKQOEMALQLOKATKQKILTYINK 120
Db 114 LOSNLSGGSTLTQOLIKLTYFSTSTSDQTSIRKQOEMALQLOKATKQKILTYINK 173
OY 121 VYMSNGNYGMOTAAQNYGKDLNNLSLPQALLAGMPQAPNOYDPSHPEAQRRLYL 180
Db 174 VYMSNGNYGMOTAAQNYGKDLNNLSLPQALLAGMPQAPNOYDPSHPEAQRRLYL 233
OY 181 SEMKNGYISABOYERAVNPTITDGLQSLKSANYPAYMDNYLKEVINQVEERTYNLT 240
Db 234 SEMKNGYISABOYERAVNPTITDGLQSLKSANYPAYMDNYLKEVINQVEERTYNLT 293
OY 241 TGMDDYTVNDQAKHMDIYNTDEVAVYPDELOVASTIVYVNSGKYIAQGAHQSSN 300
Db 294 TGMDDYTVNDQAKHMDIYNTDEVAVYPDELOVASTIVYVNSGKYIAQGAHQSSN 353
OY 301 VSEGTNOAVETRDGSTMKPTIDYAPALEYGVYSTATIVHDEPYNPGTNTPYNNMR 360
Db 354 VSEGTNOAVETRDGSTMKPTIDYAPALEYGVYSTATIVHDEPYNPGTNTPYNNMR 413
OY 361 GYFGNITLQYALQOQRNRYAVETLNKVGIDRAKTEFLNGIDIDPSIHVANAISNTTESD 420
Db 414 AFGNITTLQYALQOQRNRYAVETLNKVGIDRAKTEFLNGIDIDPSMHVANAISNTTESN 473
OY 421 KKYGASSEKMAAAYAFANGGTYTYPMTYHKVYFSDGSEKESNNGTRAMKETTAYMMTD 480
Db 474 KQYGASSEKMAAAYAFANGGTYTYPMTYHKVYFSDGSEKESNNGTRAMKETTAYMMTE 533
OY 481 MKKTVLTGTGNAYLAMLPOAGKTGTSNYTDEEIEHNHKTISOFAVAPDELFAGYTRKYSM 540
Db 534 MKKTVLTGTGNAYLAMLPOAGKTGTSNYTDEEIEHNHKTISOFAVAPDELFAGYTRKYSM 593
OY 541 AVWTGYSNRLPPLVGNGLTVAAKYRSMYTYLSEGSNPEDMNIPGLYRNGEVEFKNGAR 600
Db 594 AVWTGYSNRLPPLVGNGLTVAAKYRSMYTYLSEGSNPEDMNIPGLYRNGEVEFKNGAR 653
OY 601 STWNSPAPQOQPESTSSSSSSSTSSSTSPSTNNSTTTPNNNTQOOSNTTPOQONON 660
Db 654 STWNSPAPQOQPESTSSSSSSSTSSSTSPSTNNSTTTPNNNTQOOSNTTPOQONON 713
OY 661 POPAP 666
Db 714 POPAP 719

RESULT 13
O54948 PRELIMINARY; PRT: 719 AA.
AC 054948:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
```

OX NCBI\_TaxID=1313;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=670;  
 RX MEDLINE=93010977; Pubmed-1396576;  
 RA Martin C., Sibold C., Hakenbeck R.;  
 RT "Relatedness of penicillin-binding protein 1a genes from different  
 RT clones of penicillin-resistant Streptococcus pneumoniae isolated in  
 RT South Africa and Spain.";  
 RL EMBL J. 11:3831-3836(1992).  
 DR EMBL: X67867; CAA48067.1;  
 DR InterPro: IPR001264; Transglycosyl.  
 DR InterPro: IPR001460; Transpeptidase.  
 DR Pfam: PF00912; Transglycosyl. 1.  
 DR Pfam: PF00905; Transpeptidase; 1.  
 DR Prodom: PD001895; Transglycosyl. 1.  
 FT NON\_TER 719 719  
 SQ SEQUENCE 719 AA; 79611 MW; 1DBFDA93704930B CRC64;

Query Match 93.0%; Score 3241; DB 2; Length 719;  
 Best Local Similarity 91.4%; Pred. No. 3, 4e-168;  
 Matches 612; Conservative 36; Mismatches 18; Indels 0; Gaps 0;

OY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHFRFDHNGIDITRILGAFLRN 60  
 DB KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHFRFDHNGIDISIRILGAFLRN 113  
 OY 61 LOSNSLGGSGSTLTQOLIKLTFFSTSDQTSRKAQEWMLAIOLEKATKQELITYYINK 120  
 DB LOSNSLGGSGSTLTQOLIKLTFFSTSDQTSRKAQEWMLAIOLEKATKQELITYYINK 173  
 OY 121 VYMSNGNYGMQTAQONYCYGDLNNLSLPOLALLAGMPQAPNOYDPYSHPAADRRRLYL 180  
 DB VYMSNGNYGMQTAQONYCYGDLNNLSLPOLALLAGMPQAPNOYDPYSHPAADRRRLYL 233  
 OY 174 VYMSNGNYGMQTAQONYCYGDLNNLSLPOLALLAGMPQAPNOYDPYSHPAADRRRLYL 233  
 OY 181 SEMKNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 240  
 DB SEMKNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 293  
 OY 234 SEMKNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 293  
 OY 241 TGMADVTTNDQEAQKHLMDIYNTDEVYAYPDDELQVASTIVDSNGKVIAQLGARHOSN 300  
 DB TGMADVTTNDQEAQKHLMDIYNTDEVYAYPDDELQVASTIVDSNGKVIAQLGARHOSN 353  
 OY 294 TGMADVTTNDQEAQKHLMDIYNTDEVYAYPDDELQVASTIVDSNGKVIAQLGARHOSN 353  
 OY 301 VSRGYNQAVETNMDGSTMKPTIDVAPALEYGYVDSATVHDEPNYPGTNTPVYWMDR 360  
 DB VSRGYNQAVETNMDGSTMKPTIDVAPALEYGYVDSATVHDEPNYPGTNTPVYWMDR 413  
 OY 354 VSRGYNQAVETNMDGSTMKPTIDVAPALEYGYVDSATVHDEPNYPGTNTPVYWMDR 413  
 OY 361 GYFGNITLQYALQOQRNVPAVETLNKVGILNRAKTFNLGLIDYPSIHYNAISSNTTESD 420  
 DB AYFGNITLQYALQOQRNVPAVETLNKVGILNRAKTFNLGLIDYPSIHYNAISSNTTESD 473  
 OY 421 KRYGASSEKMAAAYAAAFANGGTYKPMYIHKVYVSDGSEKESFNVGTRAKETAYAMMD 480  
 DB KRYGASSEKMAAAYAAAFANGGTYKPMYIHKVYVSDGSEKESFNVGTRAKETAYAMMD 533  
 OY 474 KRYGASSEKMAAAYAAAFANGGTYKPMYIHKVYVSDGSEKESFNVGTRAKETAYAMMD 533  
 OY 481 MKTIVLTGTRGAAVLAAMLPOAGKTGTSNTDEIEHNIKTSQFVAPDELFAGYTRKYSM 540  
 DB MKTIVLTGTRGAAVLAAMLPOAGKTGTSNTDEIEHNIKTSQFVAPDELFAGYTRKYSM 593  
 OY 534 MKTIVLTGTRGAAVLAAMLPOAGKTGTSNTDEIEHNIKTSQFVAPDELFAGYTRKYSM 593  
 OY 541 AVWTGYSNRLTPLVNGLTVAARVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFNKNGAR 600  
 DB AVWTGYSNRLTPLVNGLTVAARVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFNKNGAR 653  
 OY 601 STWNSPAPQOPSTESSSSSDSTSSSTPSTNNSTTTNNNNNTQOQNTTPDOONON 660  
 DB STWNSPAPQOPSTESSSSSDSTSSSTPSTNNSTTTNNNNNTQOQNTTPDOONON 713  
 OY 654 STWNSPAPQOPSTESSSSSDSTSSSTPSTNNSTTTNNNNNTQOQNTTPDOONON 713  
 OY 661 POPAP 666  
 DB POPAP 719

RESULT 14  
 ID 09RET5 PRELIMINARY; PRT: 719 AA.  
 AC 09RET5;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE PENICILLIN-BINDING PROTEIN 1A.  
 GN PBP1A.  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=SP 1470;  
 RX MEDLINE=21432820; Pubmed-11549185;  
 RA Ferion A., Berche P.;  
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst  
 RT penicillin-resistant clinical isolates of Streptococcus pneumoniae  
 RT serotype 23F from the nasopharyngeal flora of children.";  
 RT J. Med. Microbiol. 50:828-832(2001).  
 RL EMBL: AF210750; AAF17260.1;  
 DR InterPro: IPR001264; Transglycosyl.  
 DR InterPro: IPR001460; Transpeptidase.  
 DR Pfam: PF00912; Transglycosyl. 1.  
 DR Pfam: PF00905; Transpeptidase; 1.  
 DR Prodom: PD001895; Transglycosyl. 1.  
 SQ SEQUENCE 719 AA; 79734 MW; 3BE77571DBC6496 CRC64;

Query Match 92.7%; Score 3231; DB 2; Length 719;  
 Best Local Similarity 91.4%; Pred. No. 1, 2e-167;  
 Matches 609; Conservative 34; Mismatches 23; Indels 0; Gaps 0;

OY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHFRFDHNGIDITRILGAFLRN 60  
 DB KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHFRFDHNGIDISIRILGAFLRN 113  
 OY 61 LOSNSLGGSGSTLTQOLIKLTFFSTSDQTSRKAQEWMLAIOLEKATKQELITYYINK 120  
 DB LOSNSLGGSGSTLTQOLIKLTFFSTSDQTSRKAQEWMLAIOLEKATKQELITYYINK 173  
 OY 121 VYMSNGNYGMQTAQONYCYGDLNNLSLPOLALLAGMPQAPNOYDPYSHPAADRRRLYL 180  
 DB VYMSNGNYGMQTAQONYCYGDLNNLSLPOLALLAGMPQAPNOYDPYSHPAADRRRLYL 233  
 OY 174 VYMSNGNYGMQTAQONYCYGDLNNLSLPOLALLAGMPQAPNOYDPYSHPAADRRRLYL 233  
 OY 181 SEMKNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 240  
 DB SEMKNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 293  
 OY 234 SEMKNGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLLT 293  
 OY 241 TGMADVTTNDQEAQKHLMDIYNTDEVYAYPDDELQVASTIVDSNGKVIAQLGARHOSN 300  
 DB TGMADVTTNDQEAQKHLMDIYNTDEVYAYPDDELQVASTIVDSNGKVIAQLGARHOSN 353  
 OY 294 TGMADVTTNDQEAQKHLMDIYNTDEVYAYPDDELQVASTIVDSNGKVIAQLGARHOSN 353  
 OY 301 VSRGYNQAVETNMDGSTMKPTIDVAPALEYGYVDSATVHDEPNYPGTNTPVYWMDR 360  
 DB VSRGYNQAVETNMDGSTMKPTIDVAPALEYGYVDSATVHDEPNYPGTNTPVYWMDR 413  
 OY 354 VSRGYNQAVETNMDGSTMKPTIDVAPALEYGYVDSATVHDEPNYPGTNTPVYWMDR 413  
 OY 361 GYFGNITLQYALQOQRNVPAVETLNKVGILNRAKTFNLGLIDYPSIHYNAISSNTTESD 420  
 DB AYFGNITLQYALQOQRNVPAVETLNKVGILNRAKTFNLGLIDYPSIHYNAISSNTTESD 473  
 OY 421 KRYGASSEKMAAAYAAAFANGGTYKPMYIHKVYVSDGSEKESFNVGTRAKETAYAMMD 480  
 DB KRYGASSEKMAAAYAAAFANGGTYKPMYIHKVYVSDGSEKESFNVGTRAKETAYAMMD 533  
 OY 474 KRYGASSEKMAAAYAAAFANGGTYKPMYIHKVYVSDGSEKESFNVGTRAKETAYAMMD 533  
 OY 481 MKTIVLTGTRGAAVLAAMLPOAGKTGTSNTDEIEHNIKTSQFVAPDELFAGYTRKYSM 540  
 DB MKTIVLTGTRGAAVLAAMLPOAGKTGTSNTDEIEHNIKTSQFVAPDELFAGYTRKYSM 593  
 OY 534 MKTIVLTGTRGAAVLAAMLPOAGKTGTSNTDEIEHNIKTSQFVAPDELFAGYTRKYSM 593  
 OY 541 AVWTGYSNRLTPLVNGLTVAARVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFNKNGAR 600  
 DB AVWTGYSNRLTPLVNGLTVAARVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFNKNGAR 653

Db	594	AWGTCYSNNLPTPIVGGDLVAAKVKYRSMITTYLSEDTHPEDMTMPDGLFRNGEIVFVNGAR	653
Oy	601	STWNSPAPQPPSTESSSSSDSSSTSSQSSSTTPSTNNSTTTNNNTTQGSNTTPDQONON	660
Db	654	PWTPESTQSSSTAESSSSSDSSSTSSQSSSTTPSTNNSTTTNNNTTQGSNTTPDQONON	713
Oy	661	POPAPQ 666	
Db	714	POPAPQ 719	
RESULT	15		
ID	054950	PRELIMINARY:	PRT: 719 AA.
AC	054950:		
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
DE	PENICILLIN-BINDING PROTEIN_1A (FRAGMENT).		
GN	PONA.		
OS	Streptococcus pneumoniae.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=1313;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-2039;		
RC	MEDLINE-93010977; PubMed-1396576;		
RX	Martin C., Stbold C., Hakenbeck R.;		
RT	"Relatedness of penicillin-binding protein 1a genes from different		
RT	clones of penicillin-resistant Streptococcus pneumoniae isolated in		
RT	South Africa and Spain.";		
RL	EMBO J. 11:3831-3836(1992).		
DR	EMBL, X67670; CAA48070.1;		
DR	InterPro: IPR001264; Transglycosyl.		
DR	InterPro: IPR001460; Transpeptidase.		
DR	Pfam: PF00912; Transglycosyl; 1.		
DR	Pfam: PF00905; Transpeptidase; 1.		
DR	ProDom: PD001895; Transglycosyl; 1.		
FT	NON_TER		
FT	719		
FT	719		
SO	SEQUENCE	719 AA: 79874 MW: FACBFBAB303D865D4 CRC64:	

Query Match	Similarity	91.6%	Score 3192	DB 2	Length 719
Best Local	Similarity	90.4%	Pred. No. 1.6e-165		
Matches	Conservative	41	Mismatches	23	Indels
					Gaps
					0

Db	414	VYFCGIIITQVLAQDSRNVTAVELINKVGLDBRAKTEFLNGIGIDVPSMHYANAISNJTSEN	473
OY	421	KRYGASSSEKMAAATAAFANGSTYYKKPMYIKHKVYPSDSESEKFEFNVGTRAKKETTAYMTD	480
Dd	474	KOYGASSSEKMAAATAAFANGSTYYKKPMYIKHKVYPSDSESEKFEFNVGTRAKKETTAYMTD	533
OY	481	MKTYLVLTGTGRNMYLAMLPAQAGTKTGSNRYDEIEJNHIKTSOFVADDELFAAGYTRKYSM	540
Dd	534	MKMTVLADGIGIRGAYLPMLPDAQGTGTSNYTDDIEIKYIKNTGVVADDEMFVGYTRKYSM	593
OY	541	AVMYGYSNRLPLVINGLITVAARYBGMYYLSCGSPEDDMNIPESGIXRGEVFPKNGAR	600
Dd	594	AVMYGYSNRLPLVINGLITVAARYBGMYYLSCGSPEDDMNIPESGIXRGEVFPKNGAR	653
OY	601	STWNSPAPQCPSTESSSSSSSDSSSTSPSTNNSTTNNPNNNTQOQNTTPDOONON	660
Dd	654	STWNSPAPQCPSTESSSSSSSDSSSTSPSTNNSTTTPDNNNTQOQNTTPDOONON	713
OY	661	POPAPQ 666	
Dd	714	POPAPQ 719	

```
Search completed: August 11, 2002, 04:34:19
Job time: 600 sec
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2002, 19:11:54 ; Search time 3654.27 Seconds  
(without alignments)  
11447.470 Million cell updates/sec

Title: US-08-961-083-1  
1999

Perfect score: 1 TAAATCTACGACATATAA.....ATCCTCAGACAGACACCA 1999

Sequence: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No. Score Match Length DB ID

Description

1	1999	100.0	1999	6	ARI20238	ARI20238 Sequence
2	1999	100.0	10711	6	BD003825	BD003825 Polynucle
3	1999	100.0	11334	1	AE007349	AE007349 Streptoco
4	1992.6	99.7	2160	1	AF210745	AF210745 Streptoco
5	1986.2	99.4	2157	1	SPON8A	X67873 S.pneumonia
6	1983	99.2	2160	1	AF210746	AF210746 Streptoco
7	1981.4	99.1	2157	1	SPON7A	X67872 S.pneumonia
8	1981.4	99.1	2160	6	AX111315	AX111315 Sequence
9	1981.4	99.1	3378	1	STRPONA	W90527 Streptococc
10	1981.4	99.1	10313	1	AE008414	AE008414 Streptoco
11	1887	94.4	2160	1	AF210752	AF210752 Streptoco
12	1704.6	85.3	2160	1	AF139884	AF139884 Streptoco
13	1704.6	85.3	2160	1	AF139885	AF139885 Streptoco
14	1704.6	85.3	2160	1	AF139886	AF139886 Streptoco
15	1677.4	83.9	2160	1	AF139887	AF139887 Streptoco
16	1677.4	83.9	2160	1	AF139888	AF139888 Streptoco
17	1677.4	83.9	2160	1	AF139889	AF139889 Streptoco
18	1677.4	83.9	2160	1	AF139890	AF139890 Streptoco
19	1677.4	83.9	2160	1	AF159448	AF159448 Streptoco
20	1677.4	83.9	2160	6	AX111316	AX111316 Sequence
21	1645.4	82.3	2157	1	SPONNA4	X67869 S.pneumonia
22	1643.8	82.2	2157	1	SPONNA3	X67868 S.pneumonia
23	1643.8	82.2	2160	1	AF210747	AF210747 Streptoco
24	1643.8	82.2	30457	2	SPNEU1914	AF449936 Streptoco
25	1642.2	82.2	2160	1	AF139883	AF139883 Streptoco
26	1642.2	82.2	2160	6	AX110436	AX110436 Sequence
27	1640.6	82.1	2160	1	AF210750	AF210750 Streptoco
28	1639	82.0	2160	1	AF210751	AF210751 Streptoco
29	1621.4	81.1	2172	1	SPCS1091A	AX111314 Sequence
30	1621.4	81.1	2172	1	SPCS1111A	249094 S.pneumonia
31	1621.4	81.1	2172	1	AF210748	249095 S.pneumonia
32	1608.6	80.5	2160	1	AF210749	AF210749 Streptoco
33	1607	80.4	2157	1	SPONNA1	X67870 S.pneumonia
34	1607	80.4	2157	1	SPONNA2	AX111327 Sequence
35	1607	80.4	2157	6	AX111407	AB006874 Streptoco
36	1563.8	78.2	2154	1	AF210748	AB006868 Streptoco
37	1556.2	77.8	1826	1	SPONNA6	AJ295852 Streptoco
38	1442.2	72.1	2157	1	SPONNA5	AX110281 Sequence
39	1247.4	62.4	1249	6	AX111327	W90528 Streptococc
40	1185.4	59.3	1195	1	AB006874	AB006879 Streptoco
41	1182.2	59.1	1195	1	AB006868	AB006874 Streptoco
42	1161.4	58.1	2184	1	SMI295852	AB006868 Streptoco
43	1122.2	56.1	1199	6	AX110281	AX110281 Sequence
44	1107.8	55.4	2528	1	STRPONA	W90528 Streptococc
45	1105.4	55.3	1195	1	AB006879	AB006879 Streptoco

## ALIGNMENTS

RESULT 1  
ARI20238 ARI20238 1999 bp DNA linear PAT 16-MAY-2001

LOCUS Sequence 1 from patent US 6159469.  
DEFINITION ARI20238  
ACCESSION ARI20238  
VERSION ARI20238.1 GI:14103814

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1999)  
AUTHORS Choi,G.H., Kunsch,C.A., Barash,S.C., Dillon,P.J., Dougherty,B.,  
Fannon,M.R. and Rosen,C.A.

TITLE Streptococcus pneumoniae antigens and vaccines  
JOURNAL Patent: US 6159469-A 12-DEC-2000;  
FEATURES location/Qualifiers  
source 1. 1999

BASE COUNT 656 a 485 c 376 g 482 t  
ORIGIN /organism="unknown"

Query Match 100.0%; Score 1999; DB 6; Length 1999;

Best Local Similarity	100.0%	Pred. No. 0;
Matches 1999; Conservative	0;	Mismatches 0;
		Indels 0;
		Gaps 0;

OY	1	TAAATCTACGCAAT	AAAATCAATCAATCTGGCTAGCTTGGGTTCTGAAACCGCGCTCA	60
Db	1	TAAATCTACGCAAT	AAAAATCAACTCATTTGCTAGCTTGGGTTCTGAAACCGCGCTCA	60
OY	61	TGCCCCAAGCTAATGATAT	TATTCACAGATTTTGTTAAGGCAATCGTTTCTATCGAAGACA	120
Db	61	TGCCCAAGCTAATGATAT	TATTCACAGATTTTGTTAAGGCAATCGTTTCTATCGAAGACA	120
OY	121	TGCGTTCTTTCGACCA	CAGGGGGATTTGATACCATCCGATATCTTGGAGACTTCTTGGCGAA	180
Db	121	TGCGTTCTTTCGACCA	CAGGGGGATTTGATACCATCCGATATCTTGGAGACTTCTTGGCGAA	180
OY	181	TCTGCAAGCAAT	TCCCTCCAAAGGTGATCAACTCTACCCCAACAGTTGATTAAAGTTGAC	240
Db	181	TCTGCAAGCAAT	TCCCTCCAAAGGTGATCAACTCTACCCCAACAGTTGATTAAAGTTGAC	240
OY	241	TTACTTTTCAACTTCG	ACTTCGACCAAGACTATTTCTCGTAAGGCTTCAGGAAGCTTGGTT	300
Db	241	TTACTTTTCAACTTCG	ACTTCGACCAAGACTATTTCTCGTAAGGCTTCAGGAAGCTTGGTT	300
OY	301	AGCGATTACAGTTA	GAACAAAAGAACCAACCAAGAAATCTTGACCTACTATATAATAA	360
Db	301	AGCGATTACAGTTA	GAACAAAAGAACCAACCAAGAAATCTTGACCTACTATATAATAA	360
OY	361	GGTCTACATGTCTAAT	TGGGAAGCTATGGAATGACAGACAGAGCTCAAAACACTATATGGTAA	420
Db	361	GGTCTACATGTCTAAT	TGGGAAGCTATGGAATGACAGACAGAGCTCAAAACACTATATGGTAA	420
OY	421	AGACCTCAATAA	TATTAAGTTTACCTCAGTTAGCCCTTCGTGGCTGGAATGCGTCAAGCAC	480
Db	421	AGACCTCAATAA	TATTAAGTTTACCTCAGTTAGCCCTTCGTGGCTGGAATGCGTCAAGCAC	480
OY	481	AAACCAATATGACCC	CTATTACATCCAGAAGACGCCAAGACCGCGGAACCTTGGTCTT	540
Db	481	AAACCAATATGACCC	CTATTACATCCAGAAGACGCCAAGACCGCGGAACCTTGGTCTT	540
OY	541	ATCTGAATGAAAA	ATCAAGGCTACATCTCGCTGCGAGACAGTATGAGAAAGCACTCAATAC	600
Db	541	ATCTGAATGAAAA	ATCAAGGCTACATCTCGCTGCGAGACAGTATGAGAAAGCACTCAATAC	600
OY	601	ACCAATTAAGT	AGACTACAAATGCTCAAAATGCAATGCAAGTAATTAATACCTGCTTACATGA	660
Db	601	ACCAATTAAGT	AGACTACAAATGCTCAAAATGCAATGCAAGTAATTAATACCTGCTTACATGA	660
OY	661	TAAATTACCTCA	AGGAAGTCATCATTCAAAGTTGAAGAAGAAACAGGCTATTAACCTACTCAC	720
Db	661	TAAATTACCTCA	AGGAAGTCATCATTCAAAGTTGAAGAAGAAACAGGCTATTAACCTACTCAC	720
OY	721	AACTGGATGATG	TCACAAATGTAAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
Db	721	AACTGGATGATG	TCACAAATGTAAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
OY	781	TTACAAATACAGA	CAATCGTTGCTTCCAGAGATGAATTCGAAGTGCCTTCTACAT	840
Db	781	TTACAAATACAGA	CAATCGTTGCTTCCAGAGATGAATTCGAAGTGCCTTCTACAT	840
OY	841	TGTTGATGTCT	TAACGTTAAAGTCATTTGCCAGCTATGAGACAGCGCATCAGTCAAGTAA	900
Db	841	TGTTGATGTCT	TAACGTTAAAGTCATTTGCCAGCTATGAGACAGCGCATCAGTCAAGTAA	900
OY	901	TGTTTCCTTGG	AATTAACCAAGAGATGAGAAACAAACCGGAGCTGGGGGATCAACTATGAA	960
Db	901	TGTTTCCTTGG	AATTAACCAAGAGATGAGAAACAAACCGGAGCTGGGGGATCAACTATGAA	960
OY	961	ACCGATACAGACT	ATATGCTCTGCTTGGAGTACGATGTCTACAGATTCAACTGCTACTAT	1020
Db	961	ACCGATACAGACT	ATATGCTCTGCTTGGAGTACGATGTCTACAGATTCAACTGCTACTAT	1020
OY	1021	CGTTACAGATGAGCC	CTATTAATCAATCCCTGGGACAAATACTCTGTTTATTAATCTGGGATAG	1080

Db	1021	GGTTCACGATGAGCCCTTA	ACTACCTCGGAGCAAAAT	ACTCCTGTTTAACTGGGATAG	1080
Qy	1081	GGGCTACTT	TGGCAATACCTTGCAAT	AGCCCTCGACAACTGGCGAAACGTCCAGC	1140
Db	1081	GGGCTACTT	TGGCAATACCTTGCAAT	AGCCCTCGACAACTGGCGAAACGTCCAGC	1140
Qy	1141	CGTGAAC	CTTAAACAAAGTGGAGCTCA	ACGGCGCAAGACTTTCCTTAAATGGTCTAGG	1200
Db	1141	CGTGAAC	CTTAAACAAAGTGGAGCTCA	ACGGCGCAAGACTTTCCTTAAATGGTCTAGG	1200
Qy	1201	AATGCACTACCCAA	GATATATGCACTACTACTCA	AATGCCATTTCCAACTAAACACACCGAATCAGA	1260
Db	1201	AATGCACTACCCAA	GATATATGCACTACTACTCA	AATGCCATTTCCAACTAAACACACCGAATCAGA	1260
Qy	1261	CAAAAAATAT	TGAGCAAGTAGTGTGAAA	AGATGGCTGCTTACGCTGGCTTTGCAAAATGG	1320
Db	1261	CAAAAAATAT	TGAGCAAGTAGTGTGAAA	AGATGGCTGCTTACGCTGGCTTTGCAAAATGG	1320
Qy	1321	TGGAACTTACT	TTAAACCAATGATATATCCAT	TAAGTCGCTTACCTTACCTGATGGAGTGA	1380
Db	1321	TGGAACTTACT	TTAAACCAATGATATATCCAT	TAAGTCGCTTACCTTACCTGATGGAGTGA	1380
Qy	1381	AGAGTTCTCTA	ATGTCGGAACCTGTCGCA	TGAAGAAACGACAGCCTATATGATGACCGA	1440
Db	1381	AGAGTTCTCTA	ATGTCGGAACCTGTCGCA	TGAAGAAACGACAGCCTATATGATGACCGA	1440
Qy	1441	CATGATGAA	AAACAGTCTTGACTTATGAA	ACTGGAGAAATGGCTATCTTGCTTGGCTCCC	1500
Db	1441	CATGATGAA	AAACAGTCTTGACTTATGAA	ACTGGAGAAATGGCTATCTTGCTTGGCTCCC	1500
Qy	1501	TCAGACTGGT	TAAACAGGAACCTCTA	ACATATACAGACGAGAAATTTGAAACACATCA	1560
Db	1501	TCAGACTGGT	TAAACAGGAACCTCTA	ACATATACAGACGAGAAATTTGAAACACATCA	1560
Qy	1561	GACCTCTCA	ATTGTTAGCACTGATGA	ACTATTTGCTGGCTATACGCTAAATATTCAT	1620
Db	1561	GACCTCTCA	ATTGTTAGCACTGATGA	ACTATTTGCTGGCTATACGCTAAATATTCAT	1620
Qy	1621	GGCTGAT	TAGGACAGGCTATCTT	CAACCGTCTGACACCACTTGTAGGCAATGGCCTTACGGT	1680
Db	1621	GGCTGAT	TAGGACAGGCTATCTT	CAACCGTCTGACACCACTTGTAGGCAATGGCCTTACGGT	1680
Qy	1681	CGCTCC	CAAAAGTTTACCGCTCT	ATGATGACCTACTGTCTGAAAGGAGCAATCCAGAGA	1740
Db	1681	CGCTCC	CAAAAGTTTACCGCTCT	ATGATGACCTACTGTCTGAAAGGAGCAATCCAGAGA	1740
Qy	1741	TTGGAAT	TATACAGAGGGGCTCT	ACAGAAATGGAGAAATTCGTATTTAAAAATGTTGCTCG	1800
Db	1741	TTGGAAT	TATACAGAGGGGCTCT	ACAGAAATGGAGAAATTCGTATTTAAAAATGTTGCTCG	1800
Qy	1801	TTTCTAG	TGGGAACCTACCGCT	CCGTCGCAACACCCCATCACTGAAAGTTCAAAGTCTATC	1860
Db	1801	TTTCTAG	TGGGAACCTACCGCT	CCGTCGTCGCAACACCCCATCACTGAAAGTTCAAAGTCTATC	1860
Qy	1861	ATCAGAT	AGTTCAACTTACAGCT	TAGCTCAACCACTCCAAACACAAATATATAGTAGAC	1920
Db	1861	ATCAGAT	AGTTCAACTTACAGCT	TAGCTCAACCACTCCAAACACAAATATATAGTAGAC	1920
Qy	1921	TACCAAT	CTCTAACTAATATAGCA	CAACATCAAAATCCCTGATCAACAATAATCAGAA	1980
Db	1921	TACCAAT	CTCTAACTAATATAGCA	CAACATCAAAATCCCTGATCAACAATAATCAGAA	1980
Qy	1981	TTCTCA	ACCAAGCAACCA	1999	
Db	1981	TTCTCA	ACCAAGCAACCA	1999	

RESULT	2			
BD003825/c				
LOCUS	BD003825	10711 bp	DNA	linear
DEFINITION	Polynucleotide of <i>Streptococcus pneumoniae</i> and sequence.			
ACCESSION	BD003825			
VERSION	BD003825.1	GI:18631786		

KEYWORDS JP 2001501833-A/145.  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 10711)  
AUTHORS Kunsch,C.A., Choi,G.H., Dillon,P.J., Rosen,C.A., Bara,S.C.,  
Fannon,M. and Dougherty,B.A.  
TITLE Polynucleotide of Streptococcus pneumoniae and sequence  
JOURNAL Patent: JP 2001501833-A 145 13-FEB-2001;  
HUMAN GENOME SCIENCES INC

COMMENT  
OS Unidentified  
PN JP 2001501833-A/145  
PD 13-FEB-2001  
PF 30-OCT-1997 JP 1998520718  
PR 31-OCT-1996 US 60/029960  
PI CHARLES A KUNSCH,GIL H CHOI,PATRICK J DILLON,CRAIG A ROSEN, PI  
STEVEN C BARASH,  
PI MICHAEL FANNON,BRIAN A DOUGHERTY  
PC C12N15/09,A01K67/027,C07K14/315,C07K16/12,C12N1/15,C12N1/19,  
PC C12N1/21,  
PC C12N5/10,C12P21/02,C12Q1/68,G06F17/30,C12N15/00,C12N5/00, PC  
G06F15/40  
CC Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers  
FT source 1..10711 /organism='unidentified'.  
FEATURES  
source Location/Qualifiers  
1..10711 /organism='unidentified'  
/db\_xref="taxon:32644"

BASE COUNT 3280 a 2056 c 2412 g 2963 t

ORIGIN

Query Match 100.0%; Score 1999; DB 6; Length 10711;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1999: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAATCTAGACAAATATAATCACTCATCTGCTGCTGGTTCGAAAGCCGCGCTCAA 60  
DB 9767 TAAATCTAGACAAATATAATCACTCATCTGCTGCTGGTTCGAAAGCCGCGCTCAA 9708  
QY 61 TCCCGAAGCTATGATATATCCACAGATTGGTTAAGCAATCGTTTCATATGGAACCA 120  
DB 9707 TCCCGAAGCTATGATATATCCACAGATTGGTTAAGCAATCGTTTCATATGGAACCA 9648  
QY 121 TCGCTTCTCGACACAGGGGGATGATACATCCGATCTGCGAGCTTTCTGGCGAA 180  
DB 9647 TCGCTTCTCGACACAGGGGGATGATACATCCGATCTGCGAGCTTTCTGGCGAA 9588  
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTCACCAACAGTTGATTAAGTTGAC 240  
DB 9587 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTCACCAACAGTTGATTAAGTTGAC 9528  
QY 241 TTACTTTTCACTTGAGACTCCGACCAAGCTATTTCTGTAAGGCTCAGAAAGCTTGTT 300  
DB 9527 TTACTTTTCACTTGAGACTCCGACCAAGCTATTTCTGTAAGGCTCAGAAAGCTTGTT 9468  
QY 301 AGCGATTGCTTAGAACAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360  
DB 9467 AGCGATTGCTTAGAACAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 9408  
QY 361 GGTACATATGCTTAATGGGAAGTATGGAATGAGACAGAGAGCTCAAAACTACTATGTA 420  
DB 9407 GGTACATATGCTTAATGGGAAGTATGGAATGAGACAGAGAGCTCAAAACTACTATGTA 9348  
QY 421 AGACCTCATATTTTAAGTTACCTCAGTTAGCTTGCTGGCTGGAATGCTCAGGACAC 480  
DB 9347 AGACCTCATATTTTAAGTTACCTCAGTTAGCTTGCTGGCTGGAATGCTCAGGACAC 9288  
QY 481 AAACCAATATGACCCCTATTCATCCAGACAGAGCCCAAGAGCCGCAACTGTGCTT 540  
|||||

DB 9287 AAACCAATATGACCCCTATTCATCCAGACAGCCCAAGAGCCGCAAACTTGCTT 9228  
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC 600  
DB 9227 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC 9168  
QY 601 ACCAATTACTGATGAGACTCAAAAGTCAAAATAGCAAGTAATTACCCTGCTTACATGGA 660  
DB 9167 ACCAATTACTGATGAGACTCAAAAGTCAAAATAGCAAGTAATTACCCTGCTTACATGGA 9108  
QY 661 TAATTACCTCAAGGAAGTATCAATCAAGTTGAAGAAACAGGCTATTAACCTACTACAC 720  
DB 9107 TAATTACCTCAAGGAAGTATCAATCAAGTTGAAGAAACAGGCTATTAACCTACTACAC 9048  
QY 721 AACTGGATGAGATGCTTACACAAATATAGACCAAGAAAGCTCAAAAACATCTGTGGATAT 780  
DB 9047 AACTGGATGAGATGCTTACACAAATATAGACCAAGAAAGCTCAAAAACATCTGTGGATAT 8988  
QY 781 TTACATACAGAGATPAGGTTGCTATCCAGAGATGGAATTGCAAGTGGCTTCTACAT 840  
DB 8987 TTACATACAGAGATPAGGTTGCTATCCAGAGATGGAATTGCAAGTGGCTTCTACAT 8928  
QY 841 TGTGATGTTTCTTAACGTAAGATCAATGCCCCAGTATGAGACAGCCATCACTCAAGTAA 900  
DB 8927 TGTGATGTTTCTTAACGTAAGATCAATGCCCCAGTATGAGACAGCCATCACTCAAGTAA 8868  
QY 901 TGTTCCTTCCGGAATTAACCAAGCAGTATGAAACAAACCCGAGCTGGGATCACTATGAA 960  
DB 8867 TGTTCCTTCCGGAATTAACCAAGCAGTATGAAACAAACCCGAGCTGGGATCACTATGAA 8808  
QY 961 ACCGATACAGAGTATGCTCTGCTGGAGTATGAGTGGTCTAGATTCAACTGCTACTAT 1020  
DB 8807 ACCGATACAGAGTATGCTCTGCTGGAGTATGAGTGGTCTAGATTCAACTGCTACTAT 8748  
QY 1021 CGTTACAGATGAGCCCTATTAACCTGAGACAAATATCTCTGTTTAACTAGGGATAG 1080  
DB 8747 CGTTACAGATGAGCCCTATTAACCTGAGACAAATATCTCTGTTTAACTAGGGATAG 8688  
QY 1081 GGGCTACTTTGGACACATCACTCTGCAATAGCCCTGCAACAATGCGGAAAGCTCCACG 1140  
DB 8687 GGGCTACTTTGGACACATCACTCTGCAATAGCCCTGCAACAATGCGGAAAGCTCCACG 8628  
QY 1141 CGTGAAGCTCTAAGCAAGGTGAGTCAACCGCGCAGAGCTTCTTAATGTTAGG 1200  
DB 8627 CGTGAAGCTCTAAGCAAGGTGAGTCAACCGCGCAGAGCTTCTTAATGTTAGG 8568  
QY 1201 AATCGACTACCCAGATTAATCACTCAATGCAATGCAATGCAATGCAATGCAATGCA 1260  
DB 8567 AATCGACTACCCAGATTAATCACTCAATGCAATGCAATGCAATGCAATGCAATGCA 8508  
QY 1261 CAAAAAATATGAGCAAGTATGTAAGAAAGTGGCTGCTTACGCTGCTTGGCAATG 1320  
DB 8507 CAAAAAATATGAGCAAGTATGTAAGAAAGTGGCTGCTTACGCTGCTTGGCAATG 8448  
QY 1321 TGAATCTACTATAACCAATGATATCATATAAGTCGCTTATAGTATGGAGTAAAA 1380  
DB 8447 TGAATCTACTATAACCAATGATATCATATAAGTCGCTTATAGTATGGAGTAAAA 8388  
QY 1381 AGAGTCTCTAATGTGGAAGTCTGCCATGGAAGAAAGACAGCTATATGATGACCA 1440  
DB 8387 AGAGTCTCTAATGTGGAAGTCTGCCATGGAAGAAAGACAGCTATATGATGACCA 8328  
QY 1441 CATGATGAAGAAAGCTTGTGGAATGGAAGTGGACGAAGTGCCTATCTTGGCTGCC 1500  
DB 8327 CATGATGAAGAAAGCTTGTGGAATGGAAGTGGACGAAGTGCCTATCTTGGCTGCC 8268  
QY 1501 TCAGGCTGTAAAGCAAGAAAGCTCTAATCATATACAGACAGAGAAATGGAAGAACCACTCA 1560  
DB 8267 TCAGGCTGTAAAGCAAGAAAGCTCTAATCATATACAGACAGAGAAATGGAAGAACCACTCA 8208  
QY 1561 GACCTCTCAATTTGTAGCACTGATGAACTATTTGCTGCTATACGCTTAAATATTCAT 1620  
DB 8207 GACCTCTCAATTTGTAGCACTGATGAACTATTTGCTGCTATACGCTTAAATATTCAT 8148  
|||||

Oy	1621	GGCTGTATGACAGAGCTATTTCTAACCGTCTGACACCACTTGTAGGCAATGGCCTTACGCT	1680	gene	/db_xref="taxon:170187"
					complement(114..2273)
					/gene="SP0369"
Db	8147	GGCTATATGACAGAGCTATTTCTAACCGTCTGACACCACTTGTAGGCAATGGCCTTACGCT	8088	CDS	complement(114..2273)
					/gene="SP0369"
Oy	1681	CGCTGCCAAGTTTACCGCTCTATGATGACCTACTGCTCTGAAGAGCAATCCAGAGA	1740		/codon_start=1
					/transl_table=11
Db	8087	CGCTGCCAAGTTTACCGCTCTATGATGACCTACTGCTCTGAAGAGCAATCCAGAGA	8028		/product="penicillin-binding protein 1A"
Oy	1741	TTTGAATATACCAAGGGGGCTCTACAGAAATGGAATTTCTGATTTAAATATGGTCTCG	1800		/protein_id="AAK74536.1"
					/db_xref="GI:14971838"
Db	8027	TTTGAATATACCAAGGGGGCTCTACAGAAATGGAATTTCTGATTTAAATATGGTCTCG	7968		
Oy	1801	TTTCTCGTGAAGACTACCTGCTCCACAAACCCCATCACTGAAGTTCAAGCTCATC	1860		/translation="MKNPTILRLKYLISFLSLVIAIVIGGVFFYYKAPLSL
					SKIVATTSKTYDKNKNLIDLGSERRYNAQANDIPDLVAIVISIEDHREFDRGID
Db	7967	TTTCTCGTGAAGACTACCTGCTCCACAAACCCCATCACTGAAGTTCAAGCTCATC	7908		TRILIGAFRLNLSNSLOGSTLQOULIKLYFSTSTSDISRAQEAEMALIDLEK
Oy	1861	ATCAGATAGTTCACCTTACAGCTAGCTCAACCACTCCACACAAATATATAGTACGAC	1920		ATKOEILTYINKYVMSGNGYGMOTAAQONYGKDLNLSLPOLALAGMPOAPNOYDP
					YSHAQODRNLVLSEKKNNGYISAEQYEAQVNTPTDGLQSLKSASNYPAVNDYL
Db	7907	ATCAGATAGTTCACCTTACAGCTAGCTCAACCACTCCACACAAATATATAGTACGAC	7848		KEYINQDEETGYMLTTGMVYNTNDOEAQRKHLMDIYNDEYVAYPDDELQVASTIV
Oy	1921	TACCAATCTCTAACAATATAGGCAACATCAATCAACCCCTGATCAACAATAATCAGAA	1980		DVSGKYIAQIGARHSSNVSFGINQAVETNRDMSGTRKPTTIDYAPALEYVYSTAT
					IVHDEPNYPGNTNPPVYNWDGTYFGNTLTQYALQDSRVNPAVETLNKGLRAKFTLN
Db	7847	TACCAATCTCTAACAATATAGGCAACATCAATCAACCCCTGATCAACAATAATCAGAA	7788		GLGIDYPSIHYSNALISNTEDSKKYGASSERMAAAYAAAFANGQTYKPMYIHKVVS
Oy	1981	TGCTCAACCAAGCAACCA 1999		gene	DGSEKERSNVTGRAMKETTAAMTDMMKTYTLTGTGRNAYLAMPDQAGKTGTSNTDE
					EIENHITSQFVAPDELFAGYTRKYSMAVMTGYNSRLTPLYGNGLTVAAKYRSMTY
Db	7787	TGCTCAACCAAGCAACCA 7769		CDS	LSRGSNPEDNNIPRGLYRNGEYFRKNGKARSWNPAPOQPPSTSSSSSDSSSTSSQSS
					STTSTNNSTTNNNTQOSNTTTPDOONMPQAP"
RESULT 3					complement(2270..2866)
AE007349/c					/gene="SP0370"
LOCUS	11334 bp	DNA	linear		complement(2270..2866)
DEFINITION	Streptococcus pneumoniae TIGR4	section 32 of 194	of the complete		/gene="SP0370"
ACCESSION	AE007349	AE005672			/codon_start=1
VERSION	AE007349.1	GI:14971837			/transl_table=11
KEYWORDS					/product="recombination protein u"
SOURCE					/protein_id="AAK74537.1"
ORGANISM	Streptococcus pneumoniae TIGR4.				/db_xref="GI:14971839"
	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;				/translation="MVNYPKRVSSOKRQTSLSOPKRNPNRNGSMFSEKMINATNDYLSQ
REFERENCE	1 (bases 1 to 11334)				GLAYIHKKPTPIQIVQDYPORSRAKIYEAFRQASTNPDGSGYNGYIIDEVEVETKQ
AUTHORS	Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D.,				KRALPKMKNFHPQIOMBOVLAQGGICVLLHFSQGPETIYLLPAPDLIRFHQDKGK
	Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J.,				SMPLYIREYGEIKAGAFQIPLYLVNIKEHLGKTR"
	Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D.,				2932..3459
	Umayam, L.A., White, O., Salzberg, S.L., Lewis, M.R., Radune, D.,				/gene="SP0371"
	Holtzapple, E., Khouli, H., Wolf, A.M., Uterback, T.R., Hansen, C.L.,				2932..3459
	McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Dickinson, T.,				/gene="SP0371"
	Hickey, E.K., Holt, I.E., Loftus, B.J., Yang, F., Smith, H.O.,				2932..3459
	Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and				/gene="SP0371"
	Fraser, C.M.				2932..3459
TITLE	Complete genome sequence of a virulent isolate of Streptococcus				/note="identified by Glimmer2; putative"
JOURNAL	pneumoniae				/codon_start=1
MEDLINE	science. 293 (5529), 498-506 (2001)				/transl_table=11
PUBMED	21357209				/product="conserved hypothetical protein"
REFERENCES	11463916				/protein_id="AAK74539.1"
	2 (bases 1 to 11334)				/db_xref="GI:14971841"
	Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D.,				/translation="MASIIFSAKIFPDEGFRGVYKQVVEFLDDVIKDYETVAA
	Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J.,				LKSLRQELADLKELETRKPKPSVQAEPLEAAITSSMTNPDILKRLNRLKEEVEFGKQ
	Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D.,				ILDSNDF"
	Umayam, L.A., White, O., Lewis, M.R., Radune, D., Holtzapple, E.,				4344..5501
	Khouli, H., Wolf, A.M., Uterback, T.R., Hansen, C.L., McDonald, L.A.,				/gene="SP0373"
	Feldblyum, T.V., Angiuoli, S., Gwinn, M., Hickey, E.K., Holt, I.E.,				4344..5501
	Loftus, B.J., Ujwal, M.L., Yang, F., Smith, H.O., Venter, J.C.,				/gene="SP0373"
	Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and Fraser, C.M.				/note="identified by Glimmer2; putative"
TITLE	Direct Submission				/codon_start=1
JOURNAL	Submitted (29-JUN-2001) The Institute for Genomic Research, 9712				/transl_table=11
FEATURES	Medical Center Dr, Rockville, MD 20850, USA				/product="conserved hypothetical protein"
Source	Location/Qualifiers				/protein_id="AAK74540.1"
	1. 11334				/db_xref="GI:14971842"
	/organism="Streptococcus pneumoniae TIGR4"				
	/strain="TIGR4"				



gene  
CDS  
/translation="MKKEFNLIATVAAGLEAVVQREVELGDCOVENGRVRFQGDVRA  
AIETINLIRADRIKIIIVGFPKATPEEELFOGVALDMENYILPLGAFPISTKRCVK  
SKIHNEPSVALISKIAVVKLOKHYARBEGLVPMENGPEEFIEVSLIKDVTATIDPT  
GSSLFKRGYRTEKGAPIKENMAAIIQLSMWYDKPLIDPTCGSGTFCIYAVIARK  
MAPLRSPFAEENNMISDRLEQVTEAAKVRELELDIMCDIDARVEIARANA  
QVAGVADITFFKQVODLRSDKINGVILSNPGERISDAGVTKLYAEMGOVFAPL  
KTMGFILTSDEAESEKSGQADKKRRLYNGLTKVDLYQYFGQKVRQEVK"  
5514..6908  
/gene="Sp0374"  
5514..6908  
/gene="Sp0374"  
/note="Identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAK74541.1"  
/db\_xref="GI:14971843"  
/translation="MSKKRRNRHKKEGQEPDFDEAKELTVGOAIRKNEEESGLP  
EDSLDVKYKOHDEIADKPAATROYKKEEVEVETQSIDDLIOEREAVERKESEASEY  
PSSBDILPLIDDEBQGLDPLLDENPTMTPEVEEONLSRLDQDSKKRKG  
ILTYLAIVSYICSAIYVYKQVARSKEIFTSSTTANOSVDVDFNTLYAFPTDSN  
KTKLNSQFBLSQLTKLDBKLBGSRHTLAKSKYDLSLATOIKAIODVNAQEPALY  
DGVLDITNAKASDAKFTDITKNTLELDKVLDAKSLKSOOTSTSSSSSQTSSSSS  
QASNTTSEKPSSESTSSRSSEVNMGLSAGVAVQRSARVAYNOSAIIDSSNSAW  
DFADGVLEOILATSRSGYITGDQYILERVNIYVNGVYLNKPDGYTLFTLNCKTG  
FVGNAGHADLDV"  
6984..8408  
/gene="Sp0375"  
6984..8408  
/gene="Sp0375"  
/note="Identified by match to TIGR protein family HMM  
TIGR00873"  
/codon\_start=1  
/transl\_table=11  
/product="6-phosphogluconate dehydrogenase,  
decarboxylating"  
/protein\_id="AAK74542.1"  
/db\_xref="GI:14971844"  
/translation="MTKANFCVGMAYMGRNLAINISRGYTAIVNRSKKETEDVIA  
CHPEKNVPSIDVESFYNSIEKPRIMLMVOAGGTDTTIOALPHLDKGLIDGNG  
TFKDDTIRNEELANSQINFITGVSGGEGALDPSIMPGQKEAEELVADVLEIS  
AKAPEDKPCVTYIGPDAGHYVMVHNGIEYDMOLIAESYDLMOHLIGSADMAE  
IFTEPMKGEIDSLYIETAITDLRSKDEGDGDPVVDYILDAAGKGTGKSGOSLDL  
GVPLSTIESEVFAVYISTYKEERYHASKVLPRPAFMEGRKALIKTIRATFESKT  
ISYNOGPAQLRASKENNMNIPRADISIMRDCGTTISRFLOKTTIDYNNRDADANL  
LDEYFDVYAKYQDAVDIVAAVQAGVPVTFESAATYFDSYRSADLPALVIAQRD  
YFGAHTYQRKDKEGTFHYSWDEK"  
8420..9109  
/gene="Sp0376"  
8420..9109  
/gene="Sp0376"  
/note="Sp0376"  
/codon\_start=1  
/transl\_table=11  
/product="DNA-binding response regulator"  
/protein\_id="AAK74543.1"  
/db\_xref="GI:14971845"  
/translation="MGRKILLEKERNLAHFLSELOKEQYRVDLVERGOKALSMAIO  
TDYDMLNVLNGLDMAODFAEKLSTRKPSVNIILDHWELOLELELYOVQFAVSYI  
KPVLEMLVARISAIFRGDFITDQHCSLMKAPRYRMLRIDVEHHTLYRGEMALTR  
REYDLATLMGSKKVLTRQELLESVMKYESATEETNIYDVYIRYRSKLDYVGQSYIK  
TVRCVGYTMOE"  
9208..10230  
/gene="Sp0377"  
9208..10230  
/gene="Sp0377"  
/codon\_start=1  
/transl\_table=11  
/product="choline binding protein C"  
/protein\_id="AAK74544.1"  
/db\_xref="GI:14971846"  
/translation="MKLLKRMQVVALAVFFEGGLIATNTVFANTTGRFVNDKRRYVY  
KDHKAIFYMHKIDKITYFGDIGEMVGMWYLEIPGQYRNLEPDNDPYNIGIQEYK  
YFEGDQALLQETDKQVLEAKTSENTGKYVGEQYPLSAEKRTYFENNNAVKTGMWIE

Query Match		100.0%;	Score 1999;	DB 1;	Length 11334;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1999;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
gene					
SSGERT					
10248..11246					
/gene="SP0378"					
10248..11246					
/gene="SP0378"					
/codon_start=1					
/transl_table=11					
/product="choline binding protein J"					
/protein_id="AAK74545.1"					
/db_xref="GI:14971847"					
QY	1	TAAATCTACGACATTAATAATCACTCATTTGCTGACCTGGTTCGAAGCGCGCTCAA	60	EGHWYLLKNGFGDDSYNPLPIGEVAKGWTODFHWITIDIRSKPAPWYLDASGKML	
DB	2115	TAAATCTACGACATTAATAATCACTCATTTGCTGACCTGGTTCGAAGCGCGCTCAA	2056	TDQKVNKNTYVYFGSSGMATGMKYVRKRYVYLUNKNDMDTKMOYLKNTYLRSSG	
QY	61	TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATCGTTCTATGGAAGACCA	120	AMVWGQDSSKWTYYLDPSNDMKTIGWKYVGNKMYLNSNAAMTGSGTIDGKYVNFPA	
DB	2055	TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATCGTTCTATGGAAGACCA	1996	SSGERT	
QY	121	TGCGTTCTTCGACACAGGGGATTTGATACCATTCGGTATCTGGGAGCTTTTTCGCCAA	180		
DB	1995	TGCGTTCTTCGACACAGGGGATTTGATACCATTCGGTATCTGGGAGCTTTTTCGCCAA	1936		
QY	181	TCTGCAAGCAATTCCTCCAGAGTGGATCACTGCACCCCAAGTTGATTAAGTTGAC	240		
DB	1935	TCTGCAAGCAATTCCTCCAGAGTGGATCACTGCACCCCAAGTTGATTAAGTTGAC	1876		
QY	241	TTACTTTTCAACTTGACCTTCGACAGCTATTTCTGTAAGGCTCAGGAAGCTTGTT	300		
DB	1875	TTACTTTTCAACTTGACCTTCGACAGCTATTTCTGTAAGGCTCAGGAAGCTTGTT	1816		
QY	301	AGCGATTGAGTTAGAAACAAACCAACCAAGCAAGAAATCTTGACTACTATATATAATA	360		
DB	1815	AGCGATTGAGTTAGAAACAAACCAACCAAGCAAGAAATCTTGACTACTATATATAATA	1756		
QY	361	GGTCTACATGTCTPATGGGAACCTATGGAATGAGACAGAGCTCAAACTACTATGTTAA	420		
DB	1755	GGTCTACATGTCTPATGGGAACCTATGGAATGAGACAGAGCTCAAACTACTATGTTAA	1696		
QY	421	AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTGCTGGCTGGAATGCTCAGGCAACC	480		
DB	1695	AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTGCTGGCTGGAATGCTCAGGCAACC	1636		
QY	481	AAACCAATATGACCCCTATTTACATTCACATTCAGACAGAGCCCAAGACCCCGAACTTG	540		
DB	1635	AAACCAATATGACCCCTATTTACATTCACATTCAGACAGAGCCCGAACTTGCTTT	1576		
QY	541	ATCTGAATATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAAGTCAATAC	600		
DB	1575	ATCTGAATATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAAGTCAATAC	1516		
QY	601	ACCAATTTACTGATGACTACAAAGTCTCAAAATCAGCAAGTAAATTCCTGCTTACATGGA	660		
DB	1515	ACCAATTTACTGATGACTACAAAGTCTCAAAATCAGCAAGTAAATTCCTGCTTACATGGA	1456		
QY	661	TAAATTAAGTCAAGAGATCATCAATCAAGTTGAAGAGCAAGAGCTATTAACCTACTAC	720		
DB	1455	TAAATTAAGTCAAGAGATCATCAATCAAGTTGAAGAGCAAGAGCTATTAACCTACTAC	1396		
QY	721	AACCTGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT	780		
DB	1395	AACCTGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT	1336		
QY	781	TTTACAATACAGACGAATACGTTGCTTATTCAGACAGATGAATTCGAAGTGCCTTACACAT	840		
DB	1335	TTTACAATACAGACGAATACGTTGCTTATTCAGACAGATGAATTCGAAGTGCCTTACACAT	1276		



OY	1141	CGTGGAACTCTTAACAGGTCGGACTCAACCGCCGCCAAGACTTTTCTTAATGCTTAGG	1200
Db	1299	CGTGGAACCTCTTAACAGAAGTCGGACTCAACCGCCGCCAAGACTTTTCTTAATGCTTAGG	1358
OY	1201	AATGCAGCTCCCCAAGATTCTACACTACAATAATGCAATTTCAAGTAACCAACCCGAATCAGA	1260
Db	1359	AATGCAGTACCACAAGATTCTACACTACAATAATGCAATTTCAAGTAACCAACCCGAATCAGA	1418
OY	1261	CAAAAAAATATGAGCAAGTAGTGAAAAGATGGCTGCTGCTTTAGCGTCCCTTTGGCAAATGG	1320
Db	1419	CAAAAAAATATGAGCAAGTAGTGAAAAGATGGCTGCTGCTTTAGCGTCCCTTTGGCAAATGG	1478
OY	1321	TGGAACTTACTTAATAACCAATGATATATCCATAAAGTCGCTTTAGTAGTAGGAGTGAAGA	1380
Db	1479	TGGAACTTACTTAATAACCAATGATATATCCATAAAGTCGCTTTAGTAGTAGGAGTGAAGA	1538
OY	1381	AGAGATTCTTAATGTGGGAAGCTGTGCCATGAAGGAAGCAACAGCCATATATGATAGCCGA	1440
Db	1539	AGAGTTTCTTAATGTGGGAAGCTGTGCCATGAAGGAAGCAACAGCCATATATGATAGCCGA	1598
OY	1441	CATGATGAAACAGCTTGTACCTTATGGAAGTGAAGCAAAATGCTTAATCTTGGCTGCC	1500
Db	1599	CATGATGAAACAGCTTGTACCTTATGGAAGTGAAGCAAAATGCTTAATCTTGGCTGCC	1658
OY	1501	TNAGCGTGGTAAACAGGAACCTCTACTATACAGAGSAGAAATTTGAAAACCATACAA	1560
Db	1659	TNAGCGTGGTAAACAGGAACCTCTACTATACAGAGSAGAAATTTGAAAACCATACAA	1718
OY	1561	GACCTCTCAATTTGTATGACACTGATGACATATTTGGTGGCTATACCGCTAAATATTCAT	1620
Db	1719	GACCTCTCAATTTGTATGACACTGATGACATATTTGGTGGCTATACCGCTAAATATTCAT	1778
OY	1621	GGCTGTATGAGACAGGCTATTTCTAACCCTGTGACACCACTTTGTAGGCATAGGCTTTACGT	1680
Db	1779	GGCTGTATGAGACAGGCTATTTCTAACCCTGTGACACCACTTTGTAGGCATAGGCTTTACGT	1838
OY	1681	CGCTGCGCAAGTTTACCCTCTATGATGACACTACCTACCTGTGGAAGGAGCAATCCAGAA	1740
Db	1839	CGCTGCGCAAGTTTACCCTCTATGATGACACTACCTGTGGAAGGAGCAATCCAGAA	1898
OY	1741	TTGGAATATACAGAGGGGCTACAGAAATGAGAAATTTGTAATTTAAAAATGGTGGCTCG	1800
Db	1899	TTGGAATATACAGAGGGGCTACAGAAATGAGAAATTTGTAATTTAAAAATGGTGGCTCG	1958
OY	1801	TTCTACGTGGAACTCACCTGCTCCACAGCAACCCCATCATCACTGAAAGTTCACGCTCATC	1860
Db	1959	TTCTACGTGGAACTCACCTGCTCCACAGCAACCCCATCATCACTGAAAGTTCACACCTCATC	2018
OY	1861	ATTCAGATACCTCAACTTCACAGCTGTAGCTCAACACATCCAGACCAANAATATATAGTAGAC	1920
Db	2019	ATTCAGATACCTCAACTTCACAGCTGTAGCTCAACACATCCAGACCAANAATATATAGTAGAC	2078
OY	1921	TACCAATCTTCAACAATATATACGCAACAATCAATAATACACCCCTGATCAACAAAATCGAA	1980
Db	2079	TACCAATCTTCAACAATATATACGCAACAATCAATAATACACCCCTGATCAACAAAATCGAA	2138
OY	1981	TTCTCAACACGACACAACCA	1999
Db	2139	TTCTCAACACGACACAACCA	2157
RESULT	5		
SPONUM8			
LOCUS		2157 bp	DNA
DEFINITION	S.pneumoniae (45607) ponh gene for penicillin-binding protein 1a.	linear	BCT 11-MAY-1995
ACCESSION	X67873		
VERSION	X67873.1	GI:47419	
KEYWORDS	penicillin-binding protein 1a; ponh gene.		
SOURCE	Streptococcus pneumoniae.		
ORGANISM	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.		

AUTHORS Hakenbeck, R.  
TITLE Direct Submission  
JOURNAL Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek.  
Genetik, Ihnestr. 73, 1000 Berlin 33, FRG  
REFERENCE 2 (bases 1 to 2157)  
AUTHORS Martin, C., Sibold, C. and Hakenbeck, R.  
TITLE Relatedness of penicillin-binding protein 1a genes from different  
clones of penicillin-resistant Streptococcus pneumoniae isolated in  
South Africa and Spain  
JOURNAL EMBO J. 11 (11), 3831-3836 (1992)  
MEDLINE 93010977  
FEATURES  
Source Location/Qualifiers  
1..2157  
/organism="Streptococcus pneumoniae"  
/strain="45607"  
/db\_xref="taxon:1313"  
/clone="pckM1-pckM16"  
1..2157  
/gene="ponA"  
1..>2157  
/gene="ponA"  
/codon\_start=1  
/transl\_table=11  
/product="penicillin-binding protein 1a"  
/protein\_id="CA48073.1"  
/db\_xref="GI:47420"  
/db\_xref="SWISS-PROT:O04707"  
translation="MNKPTILRLIKYLSISPLSIVIAIVLGGVFEVYKAPSLSE  
SKVATTSKTYDKNKLIADLSRRRNNAANDPDLVAIVYISIEHRRFDHIGID  
TIRLGLAFRLNLSLOGGSTLTQOLIKLYFSTISDQTSKRAQAWALALEOK  
ATKOEILLYINKYMSNGYKQTAONNYGKDLNLSIQALALAMPAPMYNDY  
YSHFPAODRNLVLSEKMNQYISAEQYKAVMTPIIDGLOSLSKSNYPAYMNYL  
KEVNOVEEYGNLITGMDEVYNDOEAKHLMDIINTDEYVAYPDELOVSTIV  
DVNSGVYIAOLGARHOSNSFQINQAVETNRDMGKSPITDEYAPALEYGVNSTAT  
IVHPEYVYRPTNPPYVMDRGYFNCNITQYALQDSRVPAVELINKGLKRAITPLN  
GLGIDYPSIHYSNAISSVTFESDKRYGSSSKMAAIAFAENGTYIKPMYIHKVYES  
DGSERFSNVGTFRAMKETTAAMTEMKMTVLTYGTGRNAYLAMPQAKTGTSMYDE  
BIENHITSQVAPDELFAGYTRKYSMAVWGYNRLPLVGNGLTVAAKYRSMYI  
LSEGSNEDNMNIPGLYNGEYFVFNKNGARSTWSSPAPOQPESTESSSSSDSTSSQSS  
STTPSTNSTTNNNNNQSNTPDOONQNPQAPQ"

BASE COUNT 696 a 520 c 411 g 530 t  
ORIGIN

Query Match 99.4%; Score 1986.2; DB 1; Length 2157;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1991; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAATAAATCAACTGCTGACTGGTGTCTGAACGCCGCGTCAA 60  
DB 159 TAAATCTACGACAATAAATCAACTGCTGACTGGTGTCTGAACGCCGCGTCAA 218  
QY 61 TGCCCAAGCTATGATATTTCCACAGATTGGTTAAGCAATCGTTTCTATGGAAGACA 120  
DB 219 TGCCCAAGCTATGATATTTCCACAGATTGGTTAAGCAATCGTTTCTATGGAAGACA 278  
QY 121 TCGCTTTCTGACACACAGGGGAGTATGATACATCCGTAATCCGGAAGCTTTCTGCGCAA 180  
DB 279 TCGCTTTCTGACACACAGGGGAGTATGATACATCCGTAATCCGGAAGCTTTCTGCGCAA 338  
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCAACTCTCAACCCACAGATTGATTAGTTGAC 240  
DB 339 TCTGCAAGCAATTCCTCCCAAGGTGATCAACTCTCAACCCAGATTGATTAGTTGAC 398  
QY 241 TTACTTTCACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300  
DB 399 TTACTTTCACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 458  
QY 301 AGCGATTGATTAGACAAAAGCAACAGACAGAAATCTTGACCTACTATATATAATA 360  
DB 459 AGCGATTGATTAGACAAAAGCAACAGACAGAAATCTTGACCTACTATATATAATA 518  
QY 361 GGTACATGTCTAATGGGAAGTATGAATGACAGCAGCTCAAACTACTATGTTAA 420

DB 519 GGTACATGTCTAATGGGAAGTATGAATGACAGCAGCTCAAACTACTATGTTAA 578  
QY 421 AATCCATATATTTAAGTTTACCTACCTGAGTACCTTGGCTGGAAATCCCTCAGGACAC 480  
DB 579 AATCCATATATTTAAGTTTACCTACCTGAGTACCTTGGCTGGAAATCCCTCAGGACAC 638  
QY 481 AATCCATATATGACCCCTATTTACATCCAGAACACCCCAAGACCCCAAGCTTGGCTT 540  
DB 639 AATCCATATATGACCCCTATTTACATCCAGAACACCCCAAGACCCCAAGCTTGGCTT 698  
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGTAACAGTATGAAAAAGCTCAATAC 600  
DB 699 ATCTGAATGAAAAATCAAGGCTACATCTCTGTAACAGTATGAAAAAGCTCAATAC 758  
QY 601 ACCAATTCOTGATGACATCAACAAAGTCTCAAAATCAGCAAGTATTCCTGCTTACATGA 660  
DB 759 ACCAATTCOTGATGACATCAACAAAGTCTCAAAATCAGCAAGTATTCCTGCTTACATGA 818  
QY 661 TAATTACCTCAAGAAAGTATCAATCAAGTTGAAAGAAACAGGCTATTAACCTACAC 720  
DB 819 TAATTACCTCAAGAAAGTATCAATCAAGTTGAAAGAAACAGGCTATTAACCTACAC 878  
QY 721 AACTGGATGATGTCTACACAAATGTAGACCAGAGCTCAAAAACATCTGTGGATAT 780  
DB 879 AACTGGATGATGTCTACACAAATGTAGACCAGAGCTCAAAAACATCTGTGGATAT 938  
QY 781 TTACAATACAGACGAATACGTTGCTTATCCAGACATGAATTGCAAGTCCTTACCAT 840  
DB 939 TTACAATACAGACGAATACGTTGCTTATCCAGACATGAATTGCAAGTCCTTACCAT 998  
QY 841 TGTATATGTTTCTAAGGTAAGTCAATGATCCAGTAGAGAGACCCATCAAGTCAATGA 900  
DB 999 TGTATATGTTTCTAAGGTAAGTCAATGATCCAGTAGAGAGACCCATCAAGTCAATGA 1058  
QY 901 TGTTCCTTGGGAATTAACCAAGAGTAAAGAAACCAACCGCATGGGATCAATATGAA 960  
DB 1059 TGTTCCTTGGGAATTAACCAAGAGTAAAGAAACCAACCGCATGGGATCAATATGAA 1118  
QY 961 ACCGATACAGACTATGCTCCGCTTGGAGTAGGCTGTACAGATTCAGTACGCTACAT 1020  
DB 1119 ACCGATACAGACTATGCTCCGCTTGGAGTAGGCTGTACAGATTCAGTACGCTACAT 1178  
QY 1021 GCTTCAGATGAGCCCTATTAATCTACCTGGGACAAATCTCTGTTTAACTGGGATAG 1080  
DB 1179 GCTTCAGATGAGCCCTATTAATCTACCTGGGACAAATCTCTGTTTAACTGGGATAG 1238  
QY 1081 GGGCTACTTTGGCAACATACCTTGAATAGCGCCCTGCAACATCGCAAAAGCTCCAGC 1140  
DB 1239 GGGCTACTTTGGCAACATACCTTGAATAGCGCCCTGCAACATCGCAAAAGCTCCAGC 1298  
QY 1141 CGTGAACCTGTAACCAAGGTGCGACTACACCGGCCCAAGACTTTCCTAAATGGCTTAGG 1200  
DB 1299 CGTGAACCTGTAACCAAGGTGCGACTACACCGGCCCAAGACTTTCCTAAATGGCTTAGG 1358  
QY 1201 AATGACTACCAAGTATTAATCTACTACTCAATGCAATTCCTCAAGTAAACACACGGAATCGA 1260  
DB 1359 AATGACTACCAAGTATTAATCTACTACTCAATGCAATTCCTCAAGTAAACACACGGAATCGA 1418  
QY 1261 CAAAAAATATGAGCAAGTATGTAAGAAAGATGCGTCTTACGCTTGTGCAAAATGG 1320  
DB 1419 CAAAAAATATGAGCAAGTATGTAAGAAAGATGCGTCTTACGCTTGTGCAAAATGG 1478  
QY 1321 TGGAACTACTATAAACAATGATATATCCATTAAGTCTGTTAGTATGGAGAGTAA 1380  
DB 1479 TGGAACTACTATAAACAATGATATATCCATTAAGTCTGTTAGTATGGAGAGTAA 1538  
QY 1381 AGAGTTCTCTAATGTGCGAATCTGTCATGAAAGAAAGCAAGCTTATATGATGACCGA 1440  
DB 1539 AGAGTTCTCTAATGTGCGAATCTGTCATGAAAGAAAGCAAGCTTATATGATGACCGA 1598  
QY 1441 CATGATGAAAAAGCTTGAATATGAGATGACGAAGTCTATCTTGGCTGCC 1500



Db 819 TAATTACCTCAAGAGATCATCATCAAGTGAAGAGAACAGGCTATTAACCTGCTCAC 878  
Qy 721 AACTGGATGATGATGCTCTACACAAATGTAGACCAGAAAGCTCAAAAACATCTGTGGATAT 780  
Db 879 AACTGGATGATGATGCTCTACACAAATGTAGACCAGAAAGCTCAAAAACATCTGTGGATAT 938  
Qy 781 TTACAATACAGACGAATAGTGGCTATTCAGACGATGAATTGCAAGTGGCTTACCAT 840  
Db 939 TTACAATACAGACGAATAGTGGCTATTCAGACGATGAATTGCAAGTGGCTTACCAT 998  
Qy 841 TGTGATGTTTCTAAGCGGAAAGTCCAGCTAGAGAGACGCCATCATCAAGTAA 900  
Db 999 TGTGATGTTTCTAAGCGGAAAGTCCAGCTAGAGAGACGCCATCATCAAGTAA 1058  
Qy 901 TGTTCCTTCGGAATTTAACCAAGACGATAGAAACAAACCGCATGGGATCAACTATGAA 960  
Db 1059 TGTTCCTTCGGAATTTAACCAAGACGATAGAAACAAACCGCATGGGATCAACTATGAA 1118  
Qy 961 ACCGATACAGACTATGCTCTGCTTGAGTAGCGGTGTCTACGATTCAACTGCTACTAT 1020  
Db 1119 ACCGATACAGACTATGCTCTGCTTGAGTAGCGGTGTCTACGATTCAACTGCTACTAT 1178  
Qy 1021 CGTTACGATGAGCCCTATTAATACCTGGGACAAATACCTGTTTATTAAGTGGGATAG 1080  
Db 1179 CGTTACGATGAGCCCTATTAATACCTGGGACAAATACCTGTTTATTAAGTGGGATAG 1238  
Qy 1081 GGGCTACTTTGGCAACATCATCCTTGCAATATACGCCCTGCACAAATGCGCAACGTCACAG 1140  
Db 1239 GGGCTACTTTGGCAACATCATCCTTGCAATATACGCCCTGCACAAATGCGCAACGTCACAG 1298  
Qy 1141 CGTGAACACTTAAACAAGCTGGACATACCCGCCCAAGACTTTCTTAATGCTGTAGG 1200  
Db 1299 CGTGAACACTTAAACAAGCTGGACATACCCGCCCAAGACTTTCTTAATGCTGTAGG 1358  
Qy 1201 AATGACATACCCAGATATTTCACTACTCAAAATGCATTTTCAAGTAAACACAAACGAATCAGA 1260  
Db 1359 AATGACATACCCAGATATTTCACTACTCAAAATGCATTTTCAAGTAAACACAAACGAATCAGA 1418  
Qy 1261 CAAAAAATATGAGACAGTAGTGAAGAAAGATGCTGCTACGCTGCTTGGCAAAATGG 1320  
Db 1419 CAAAAAATATGAGACAGTAGTGAAGAAAGATGCTGCTGCTTGGCAAAATGG 1478  
Qy 1321 TGGACATTAATAAACAATGATATATCCATTAAGTCTGTTTACGATGGAGTGAATA 1380  
Db 1479 TGGACATTAATAAACAATGATATATCCATTAAGTCTGTTTACGATGGAGTGAATA 1538  
Qy 1381 AGAGTTCTCTAATGTCGGAACGTCGTCATGAAGAAACGACAGCTATATGATGACGA 1440  
Db 1539 AGAGTTCTCTAATGTCGGAACGTCGTCATGAAGAAACGACAGCTATATGATGACGA 1598  
Qy 1441 CATGATGAAGACGCTTGAATTTATGAACTGAGCAAAATGCTTATGCTTGGCTCCC 1500  
Db 1599 AATGATGAAGACGCTTGAATTTATGAACTGAGCAAAATGCTTATGCTTGGCTCCC 1658  
Qy 1501 TCAGGCTGTAAGACAGAACCTCTTAATTAAGACAGAGAAATTTGAAACCATCA 1560  
Db 1659 TCAGGCTGTAAGACAGAACCTCTTAATTAAGACAGAGAAATTTGAAACCATCA 1718  
Qy 1561 GACCTCTCAATTTGAGACCCGATGAACTATTTGCTGGCTATAGCGCAATATTTCAAT 1620  
Db 1719 GACCTCTCAATTTGAGACCCGATGAACTATTTGCTGGCTATAGCGCAATATTTCAAT 1778  
Qy 1621 GAGCTGATGAGACGCTATTTCAACCGCTGACACACACTTTGAGCAATGAGCTTACGCT 1680  
Db 1779 GAGCTGATGAGACGCTATTTCAACCGCTGACACACACTTTGAGCAATGAGCTTACGCT 1838  
Qy 1681 CGCTGCCAAAGTTTACCGCTCTATATGACCTACTGCTGTAAGAGACCAATCCAGAGA 1740  
Db 1839 CGCTGCCAAAGTTTACCGCTCTATATGACCTACTGCTGTAAGAGACCAATCCAGAGA 1898  
Qy 1741 TTGGAATATACAGAGGAGCTCTACAGAAATGGAATTTGATTTAAATAATGAGCTCG 1800  
Db 1899 TTGGAATATACAGAGGAGCTCTACAGAAATGGAATTTGATTTAAATAATGAGCTCG 1958

Qy 1801 TTCTACGTGGAACCTACCTGCTCCACAAACACCCCATCACTGAAGTTCAAGCTCATC 1860  
Db 1959 TTCTACGTGGAACCTACCTGCTCCACAAACACCCCATCACTGAAGTTCAAGCTCATC 2018  
Qy 1861 ATCAGATAGTTCACTTCACTACAGTCTAGCTCAACCACTCCACACAAATATAGTACGAC 1920  
Db 2019 ATCAGATAGTTCACTTCACTACAGTCTAGCTCAACCACTCCACACAAATATAGTACGAC 2078  
Qy 1921 TACCAATCTTACCAATATATACCAATCAATCAATCAATCAATCAATCAATCAATCA 1980  
Db 2079 TACCAATCTTACCAATATATACCAATCAATCAATCAATCAATCAATCAATCAATCA 2138  
Qy 1981 TCCTCAACCGACCAACCA 1999  
Db 2139 TCCTCAACCGACCAACCA 2157

RESULT 7  
SPONAN7  
LOCUS S.pneumoniae (63915) ponA gene for penicillin-binding protein 1a. 2157 bp DNA linear BCT 11-MAY-1995  
ACCESSION X67872.1 GI:47417  
VERSION X67872.1 GI:47417  
KEYWORDS penicillin-binding protein 1a; ponA gene.  
SOURCE Streptococcus pneumoniae.  
ORGANISM Streptococcus pneumoniae  
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.

REFERENCE 1 (bases 1 to 2157)  
AUTHORS Hakenbeck, R.  
TITLE Direct Submision  
JOURNAL Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek. Genetik, Ihnestr. 73, 1000 Berlin 33, FRG  
REFERENCE 2 (bases 1 to 2157)  
AUTHORS Martin, C., Sibold, C. and Hakenbeck, R.  
TITLE Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain  
JOURNAL EMBO J. 11 (11), 3831-3836 (1992)  
MEDLINE 93010977

FEATURES  
source location/Qualifiers

1..2157 /organism="Streptococcus pneumoniae"  
/strain="63915"  
/db\_xref="taxon:1313"  
/clone="pckM1-pckM16"  
1..2157 /gene="ponA"  
1..>2157 /gene="ponA"

1..>2157 /gene="ponA"  
/codon\_start=1  
/transl\_table=1  
/product="penicillin-binding protein 1a"  
/protein\_id="CA48072.1"  
/db\_xref="GI:47418"  
/db\_xref="Swiss-Prot:O04707"  
/translation="MNKPTILRLIKYLSIFSLVIAIIVLGGVFFYYKAPSLSE  
SKIVATNSRIYDNKKNOLIDGSERRVNAQANDIPNDVAIYISIEHREFDREGD  
TIRILGAFIRLQNSLOGSSTLDQOLIKLTFSTSDVITISKABEAWAIDLEOK  
ATKQELILYITNKYMSNGNIGMOTAAQNTYIGKDLNLSLQQLALGMPAPMDP  
YSHEPADDRNLVLSEKNOGYISAEQYKAVNPITDGLQSLKASNPAYMDY  
KEVINQVEETGYNLTLTGMDVYTNVDQEAQKHLMDIYNDQVAYAPDDELQVASTY  
DVSNGKYAIOAGARHOSNVSGINQAVETRDGSMKPIYDPALEVGIDYSTAT  
IVHDKPYNPETNPYNNWDGFGNGITLOVALQDSRPVETLNVGLBRAKTFPLN  
GLGIDYISIHNSNISSTPSSDKKYKASSSKMAAAYAAPANGCTYYKPKYIHKVFS  
DGSEKESNVGTRAIKETATYMTDMKTYLSTGTRNATLAMIPOAKGTSTNRYTE  
EIEHNIKTSQVAPDELFAGYTRKYSAWVWGYSNRLPLVNGELTYAAKYSRMNTY  
LSEGSNPDENIIPGLYRNGEFVFKNGARSTWNPAPQPPSTESSSSSSSSTSOS  
STPSTNNSTTNNNTTQGSNTTTPDOONONPQAPQ"

BASE COUNT  
ORIGIN

699 a 521 c 407 g 530 t

Query Match 99.1% Score 1981.4; DB 1; Length 2157;  
Best Local Similarity 99.4% Pred. No. 0; Mismatches 11; Indels 0; Gaps 0;  
Matches 1988; Conservative 0;

```
QY 1 TAAATCTACGACATATAAATCAATCTGCTGACTTGGGTTCTGAACGCCGGCTCA 60
  |||
Db 159 TAAATCTACGACATATAAATCAATCTGCTGACTTGGGTTCTGAACGCCGGCTCA 218
QY 61 TGCACAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACA 120
  |||
Db 219 TGCCCAACCTAATGATATTTCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGACA 278
QY 121 TCGCTCTTGACACACAGGGGGATTGATACCATCCGATCCTCGGAGCTTTCTGCGAA 180
  |||
Db 279 TCGCTCTTGACACACAGGGGGATTGATACCATCCGATCCTCGGAGCTTTCTGCGAA 338
QY 181 TCTGCAAGCAATTCCTCCACAGGTGATCACTCTACCCACAGTTGATTAAATTGAC 240
  |||
Db 339 TCTGCAAGCAATTCCTCCACAGGTGATCACTCTACCTCAACAGTTGATTAAATTGAC 398
QY 241 TTACTTTCAACTTCGACTTCGACAGCAATTTCTGTAAGGCTCAGGAAGCTTGGTT 300
  |||
Db 399 TTACTTTCAACTTCGACTTCGACAGCAATTTCTGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCATTCAGTTAGAACAAAAGCAACCAACAGCAAAATCTTGACTCTATATAATAA 360
  |||
Db 459 AGCATTCAGTTAGAACAAAAGCAACCAACAGCAAAATCTTGACTCTATATAATAA 518
QY 361 GGTCTACATGCTAATGAGCACTATGAAATGACAGACAGCTCAAACTCTATGTAA 420
  |||
Db 519 GGTCTACATGCTAATGAGCACTATGAAATGACAGACAGCTCAAACTCTATGTAA 578
QY 421 AGACCTCAATATTTAAGTTTAACTTACCTAGCTTGGCTGCTGGAATGCTCAGGACC 480
  |||
Db 579 AGACCTCAATATTTAAGTTTAACTTACCTAGCTTGGCTGGAATGCTCAGGACC 638
QY 481 AAACCAATATGACCCCTTACATCCAGAGCAAGCCCAACCCGCAAACTTGTGCTT 540
  |||
Db 639 AAACCAATATATCCCTTATACATCCAGAGCAAGCCCAACCCGCAAACTTGTGCTT 698
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTGCTGAACAGATGAGAAAGCAATATAC 600
  |||
Db 699 ATCTGAATGAAAAATCAAGGCTACATCTGCTGAACAGATGAGAAAGCAATATAC 758
QY 601 ACCAATTTACTGATGACTACAAAGTCTCAATTCAGCAAGTAATTAACCTGCTTACATGA 660
  |||
Db 759 ACCAATTTACTGATGACTACAAAGTCTCAATTCAGCAAGTAATTAACCTGCTTACATGA 818
QY 661 TAAATCTCAAGAGATCATCAAGTTGAGAGAAACAGGCTTAACTACTAC 720
  |||
Db 819 TAAATCTCAAGAGATCATCAAGTTGAGAGAAACAGGCTTAACTACTAC 878
QY 721 AACTGGATGATGCTCTCACAAAATGTAGACCAAGAACTCAAAAACATCTGTGGGATAT 780
  |||
Db 879 AACTGGATGATGCTCTCACAAAATGTAGACCAAGAACTCAAAAACATCTGTGGGATAT 938
QY 781 TTACAATACAGAGCAATAGTTGCTTATCCAGAGATGAATTTGCAAGTCCGCTTACCAT 840
  |||
Db 939 TTACAATACAGAGCAATAGTTGCTTATCCAGAGATGAATTTGCAAGTCCGCTTACCAT 998
QY 841 TGTGTGATGTTCTTAACGGTAAAGTATGCGCAGCTAGAGACAGCCATCAAGTAA 900
  |||
Db 999 TGTGTGATGTTCTTAACGGTAAAGTATGCGCAGCTAGAGACAGCCATCAAGTAA 1058
QY 901 TGTTCCTTGGAAATTAACCAAGCAGTAGAAACAAACCGCAGCTGGGGATCAACATGAA 960
  |||
Db 1059 TGTTCCTTGGAAATTAACCAAGCAGTAGAAACAAACCGCAGCTGGGGATCAACATGAA 1118
QY 961 ACCGATACAGACTATGCTCTGCTTGGAGTACGGTGTACAGATTCAACTGCTACTAT 1020
  |||
Db 1119 ACCGATACAGACTATGCTCTGCTTGGAGTACGGTGTACAGATTCAACTGCTACTAT 1178
```

```
QY 1021 CGTTCAGATGAGCCCTTAATACCTGGAGCAATACTCCTGTTTAACTGGGATAG 1080
  |||
Db 1179 CGTTCAGATGAGCCCTTAATACCTGGAGCAATACTCCTGTTTAACTGGGATAG 1238
QY 1081 GGGCTACTTTGGCAATCACACCTTGGCAATAGCCCTGGCAACATAGCGCAACGTECCAGC 1140
  |||
Db 1239 GGGCTACTTTGGCAATCACACCTTGGCAATAGCCCTGGCAACATAGCGCAACGTECCAGC 1298
QY 1141 CGTGAACCTCTAAACAAAGTGGAGCTCAACCGCGCCCAAGACTTTCCTAATGTCGTAGG 1200
  |||
Db 1299 CGTGAACCTCTAAACAAAGTGGAGCTCAACCGCGCCCAAGACTTTCCTAATGTCGTAGG 1358
QY 1201 AATGCACTAACCAATATCTACTCAAAATGCCATTCAGTAACAAACCGAATGCA 1260
  |||
Db 1359 AATGCACTAACCAATATCTACTCAAAATGCCATTCAGTAACAAACCGAATGCA 1418
QY 1261 CAAAAAATATGAGCAATAGTAAGAAATAGGCTCTGCTTACGCTGCTTTCGCAATAGG 1320
  |||
Db 1419 CAAAAAATATGAGCAATAGTAAGAAATAGGCTCTGCTTACGCTTACGCTTTCGCAATAGG 1478
QY 1321 TGGAACTTACTATAACCAATGTATATCCATTAAGTGTCTTTAGTGAATGGAGTGAATA 1380
  |||
Db 1479 TGGAACTTACTATAACCAATGTATATCCATTAAGTGTCTTTAGTGAATGGAGTGAATA 1538
QY 1381 AGAGTCTCTAATGTGGAACCTCGTCCATGAGGAACAGAGGCTATATGATGACCGA 1440
  |||
Db 1539 AGAGTCTCTAATGTGGAACCTCGTCCATGAGGAACAGAGGCTATATGATGACCGA 1598
QY 1441 CATGATGAAAAACAGCTCTGACTTATGAACTGGAAGTAAGCTCTTCTGCTGCTCC 1500
  |||
Db 1599 CATGATGAAAAACAGCTCTGACTTATGAACTGGAAGTAAGCTCTTCTGCTGCTCC 1658
QY 1501 TCAGGCTGTAAACAGCAACCTCTAATATACAGACAGCAAGAAATGAAACCAATCA 1560
  |||
Db 1659 TCAGGCTGTAAACAGCAACCTCTAATATACAGACAGCAAGAAATGAAACCAATCA 1718
QY 1561 GACCTCTCAATTTAGACACCTGATGAACTATTTGCTGGCTATACGGTAAATATTCAT 1620
  |||
Db 1719 GACCTCTCAATTTAGACACCTGATGAACTATTTGCTGGCTATACGGTAAATATTCAT 1778
QY 1621 GCGTGTATGACAGGCTATTTAAACCGTCTGACACACCTTGTAGGCAATGGCCTTACGCT 1680
  |||
Db 1779 GCGTGTATGACAGGCTATTTAAACCGTCTGACACACCTTGTAGGCAATGGCCTTACGCT 1838
QY 1681 CGCTGCAAAAGTTTACGCTCTATGATGACCTACTGCTGTAAGGAGCAATCCAGAGA 1740
  |||
Db 1839 CGCTGCAAAAGTTTACGCTCTATGATGACCTACTGCTGTAAGGAGCAATCCAGAGA 1898
QY 1741 TTGGAATTTACCAAGGGGCTCTACAGAAATGAGAAATTCGTATTTAAATATGGCTCG 1800
  |||
Db 1899 TTGGAATTTACCAAGGGGCTCTACAGAAATGAGAAATTCGTATTTAAATATGGCTCG 1958
QY 1801 TTCTAAGTGAAGTACCTGCTCCACACAAACCCCATCACTGAAGTTCAAGCTCATC 1860
  |||
Db 1959 TTCTAAGTGAAGTACCTGCTCCACACAAACCCCATCACTGAAGTTCAAGCTCATC 2018
QY 1861 ATCAGATAGTTCACCTTACAGTGTAGCTCAACCACTCCAAAGCAAAATATATAGTACAG 1920
  |||
Db 2019 ATCAGATAGTTCACCTTACAGTGTAGCTCAACCACTCCAAAGCAAAATATATAGTACAG 2078
QY 1921 TACCAATCTTAACAATATATAGCAACATCAATATCAACCCCTGATTCACAAATACAGAA 1980
  |||
Db 2079 TACCAATCTTAACAATATATAGCAACATCAATATCAACCCCTGATTCACAAATACAGAA 2138
QY 1981 TCTCTCAACGACAGCAACCA 1999
  |||
Db 2139 TCTCTCAACGACAGCAACCA 2157
```

RESULT 8  
AX11315 2160 bp DNA linear PAT 30-APR-2001  
LOCUS  
DEFINITION Sequence 2048 from Patent WO0123604.



ACCESSION AX111315  
VERSION AX111315.1 GI:13927607  
KEYWORDS  
SOURCE Streptococcus pneumoniae.  
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
Streptococcus pneumoniae.  
REFERENCE 1 (bases 1 to 2160)  
AUTHORS Bergeron,M.G., Boltslot,M., Huletsky,A., m Nard,C., Ouellette,M.,  
Picard,F.J. and Roy,P.H.  
TITLE Highly conserved genes and their use to generate probes and primers  
for detection of microorganisms  
JOURNAL Patent: WO 0123604-A 2048 05-Apr-2001;  
Infectio Diagnostic (I.D.I.) INC. (CA)  
location/Qualifiers  
FEATURES  
source 1. 2160  
/organism="Streptococcus pneumoniae"  
/strain="R6"  
/db\_xref="taxon:1313"  
BASE COUNT 697 a 522 c 413 g 528 t  
ORIGIN  
Query Match 99.1%; Score 1981.4; DB 6; Length 2160;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 1 TAAATCTAGACAAATAAATCAATCATTTGCTGACTGGGCTCTGAAGCGCGCTCAA 60  
DB 159 TAAATCTAGACAAATAAATCAATCATTTGCTGACTGGGCTCTGAAGCGCGCTCAA 218  
QY 61 TGCCCAAGCTAATGATATTCCTCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGCA 120  
DB 219 TGCCCAAGCTAATGATATTCCTCCACAGATTTGGTTAAGCAATCGTTTCTATCGAAGCA 278  
QY 121 TCGCTTTTGACACCAAGGGGATTTGATACATCCGTAATCTGGAGCTTTCTTGGCGAA 180  
DB 279 TCGCTTTTGACACCAAGGGGATTTGATACATCCGTAATCTGGAGCTTTCTTGGCGAA 338  
QY 181 TCTGCAAGCAATTCCTCCAGGTGATCAATCTCAACCCAGCAAGTTAAGTTAGTAC 240  
DB 339 TCTGCAAGCAATTCCTCCAGGTGATCAATCTCAACCCAGCAAGTTAAGTTAGTAC 398  
QY 241 TTACTTTTCAACTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300  
DB 399 TTACTTTTCAACTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458  
QY 301 AGCGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360  
DB 459 AGCGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518  
QY 361 GGTCTACATGCTAATGGGAATATGGAATGACAGACAGCAGCTCAAAAATCTATGTGTA 420  
DB 519 GGTCTACATGCTAATGGGAATATGGAATGACAGACAGCAGCTCAAAAATCTATGTGTA 578  
QY 421 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGTGGTGGAAATGCCCTCAGGACC 480  
DB 579 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGTGGTGGAAATGCCCTCAGGACC 638  
QY 481 AAACCATATGACCCCTATTTACATCCAGAAAGCAAGCCAGACCGCGAAACTTGGTCTT 540  
DB 639 AAACCATATGACCCCTATTTACATCCAGAAAGCAAGCCAGACCGCGAAACTTGGTCTT 698  
QY 541 ATCTGAATGAAAATGAAGGCTACATCTCTGTAACAGATGAGAAAGAGCAATATAC 600  
DB 699 ATCTGAATGAAAATGAAGGCTACATCTCTGTAACAGATGAGAAAGAGCAATATAC 758  
QY 601 ACCAATTAAGTAGACTACAAAGCTCAAAATGACAAAGTAATTAATCCCTGCTCATGGA 660  
DB 759 ACCAATTAAGTAGAGGTACAAAGCTCAAAATGACAAAGTAATTAATCCCTGCTCATGGA 818  
QY 661 TAATTAACCTCAGGAAGTCAATCAATCAAGTTGAAGAAAGAAAGAGCTTAATCACTATCAC 720  
DB 720 TAATTAACCTCAGGAAGTCAATCAATCAAGTTGAAGAAAGAAAGAGCTTAATCACTATCAC 780

DB 819 TAATTAACCTCAGGAAGTCAATCAATCAAGTTGAAGAAAGAAAGAGCTTAATCACTATCAC 878  
QY 721 AACTGGATGGATGCTTACACAAATGTAGACCAAGAGCTCAAAAAGATCTGTGGGATAT 780  
DB 879 AACTGGATGGATGCTTACACAAATGTAGACCAAGAGCTCAAAAAGATCTGTGGGATAT 938  
QY 781 TTCAATATACAGCAATATGCTTGGCTATCCAGACGATGAATGCAAGTCCGTTTACCAT 840  
DB 939 TTCAATATACAGCAATATGCTTGGCTATCCAGACGATGAATGCAAGTCCGTTTACCAT 998  
QY 841 TGTGATGTTTCTTAACGGTAAAGTCAATTTGCCACCTGAGACAGCCATCAGTCAAGTAA 900  
DB 999 TGTGATGTTTCTTAACGGTAAAGTCAATTTGCCACCTGAGACAGCCATCAGTCAAGTAA 1058  
QY 901 TGTTCCTTGGGAATTAACCAAGCAGTAGAAGCAAAACCGCAGCTGGGGATCAACTATGAA 960  
DB 1059 TGTTCCTTGGGAATTAACCAAGCAGTAGAAGCAAAACCGCAGCTGGGGATCAACTATGAA 1118  
QY 961 ACCGATCAGACGATATGCTCTGCTTGGAGTACGGGTCTTACATTCATCTGCTACTAT 1020  
DB 1119 ACCGATCAGACGATATGCTCTGCTTGGAGTACGGGTCTTACATTCATCTGCTACTAT 1178  
QY 1021 CGTTCAGATGAGCCCTATATACCTCCGGGACAAATATCTGTTTATTAATCTGGATAG 1080  
DB 1179 CGTTCAGATGAGCCCTATATACCTCCGGGACAAATATCTGTTTATTAATCTGGATAG 1238  
QY 1081 GGGCTACTTTGGCAACATCACTTGCATATACGCCCTGCAACATACGGGAAACGTCGCCAGC 1140  
DB 1239 GGGCTACTTTGGCAACATCACTTGCATATACGCCCTGCAACATACGGGAAACGTCGCCAGC 1298  
QY 1141 CGTGAACATCTTAACAAGAGTCCGACTCAACCGCGCCACGCTTCTTAATGGTCTAAG 1200  
DB 1299 CGTGAACATCTTAACAAGAGTCCGACTCAACCGCGCCACGCTTCTTAATGGTCTAAG 1358  
QY 1201 AATGACATACCAAGTATTCCTACTCAATATGCCATTTCAAGTAACAAACGGAATGGA 1260  
DB 1359 AATGACATACCAAGTATTCCTACTCAATATGCCATTTCAAGTAACAAACGGAATGGA 1418  
QY 1261 CAAAAAATATGAGACCAAGTAGTGAAGAAAGATGGCTGCTTACGCTTGGCTTGGCAATGG 1320  
DB 1419 CAAAAAATATGAGACCAAGTAGTGAAGAAAGATGGCTGCTTACGCTTGGCTTGGCAATGG 1478  
QY 1321 TGAACCTTACTATTAACCAATGTAATTCATTAAGTCTTTTATGATGGAGTGAAGA 1380  
DB 1479 TGAACCTTACTATTAACCAATGTAATTCATTAAGTCTTTTATGATGGAGTGAAGA 1538  
QY 1381 AGAGTTCTCTAATGCGGAATCGGCAATGAGGAAAGCAAGACGCTTATGATGACCGA 1440  
DB 1539 AGAGTTCTCTAATGCGGAACTCGTCCATGAGAAAGAAACGACAGCTTATGATGACCGA 1598  
QY 1441 CATGATGAAAACAGTCTTGAATATGAACTGACAGAAATGCTATCTTGGCTGCC 1500  
DB 1599 CATGATGAAAACAGTCTTGAATATGAACTGACAGAAATGCTATCTTGGCTGCC 1658  
QY 1501 TCAGGCTGGTAAACAGAGACCTTAATCTATACAGAGAGAGAAATGAAAACCAATCAA 1560  
DB 1659 TCAGGCTGGTAAACAGAGACCTTAATCTATACAGAGAGAGAAATGAAAACCAATCAA 1718  
QY 1561 GACCTCTCAATTTGTAGACCTGATGAACATTTTGGTGGCTATACGGCTTAATATTCAT 1620  
DB 1719 GACCTCTCAATTTGTAGACCTGATGAACATTTTGGTGGCTATACGGCTTAATATTCAT 1778  
QY 1621 GGTGTATGACAGCTATTTCAACCGTCTGACACCACTTTGTAGGCAATGGCCTTACGGT 1680  
DB 1779 GGTGTATGACAGCTATTTCAACCGTCTGACACCACTTTGTAGGCAATGGCCTTACGGT 1838  
QY 1681 CGCTGCCAAAGTTAACGGCTCTATGATGACCTACCTGCTGAAGAGCAATCCAGAAAGA 1740  
DB 1839 CGCTGCCAAAGTTAACGGCTCTATGATGACCTACCTGCTGAAGAGCAATCCAGAAAGA 1898  
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAATATGCTATTTTAAAAATGGTGGCTG 1800  
DB 1899 TTGGAATATACAGAGGGGCTCTACAGAAATGGAATATGCTATTTTAAAAATGGTGGCTG 1958





Qy	901	TGTTTCCTTCGGAAATTAAACCAAGAGTAGAAACCAACCGGAGCTGGGGATCAACTATGAA	966
Db	2004	TGTTTCCTTCGGAAATTAAACCAAGAGTAGAAACCAACCGGAGCTGGGGATCAACTATGAA	206
Qy	961	ACCGATCAAGACATATGCTCTCGCTTGAGAGTACGGTGTCTACGATTCACATGCTCTAT	1021
Db	2064	ACCGATCAAGACATATGCTCTCGCTTGAGAGTACGGTGTCTACGATTCACATGCTCTAT	212
Qy	1021	CGTTACAGATGAGCCCTTAATACTAACCTGGGACAAATACTCGTGTATTATTAAGTGGATAG	108
Db	2124	CGTTACAGATGAGCCCTTAATACCTCCGGGACAAATACCCCTGTTTATTAAGTGGATAG	218
Qy	1081	GGGGCTACTTTGGCAACATCAACCTTGGCAATAGGCGCTTCACAACTACGCGAAAGCTCCAC	1141
Db	2184	GGGGCTACTTTGGCAACATCAACCTTGGCAATAGGCGCTTCACAACTACGCGAAAGCTCCAC	224
Qy	1141	CGTGGAAACCTTAACCAAGGTGCGACATCAACCGCGCCACAGACTTCTCTAAATGTTCTAG	1201
Db	2244	CGTGGAAACCTTAACCAAGGTGCGACATCAACCGCGCCACAGACTTCTCTAAATGTTCTAG	230
Qy	1201	AATGCACTACCCAGTATTCTACTATCTCAATGCCATTTCAGATTAACACACGAAATCAGA	1261
Db	2304	AATGCACTACCCAGTATTCTACTACTCAATATGCCATTTCAGATTAACACACGAAATCAGA	236
Qy	1261	CAAAAAATATGAGACCAAGTATGTGAAAAAGATGGGCGCGCTTACGCGCGCTTTGCCAAATG	1321
Db	2364	CAAAAAATATGAGACCAAGTATGTGAAAAAGATGGGCGCGCTTACGCGCGCTTTGCCAAATG	242
Qy	1321	TGGAACCTACTATTAACCAATGTATATCATCAATAAAGTCGTCCTTATAGTAGAGAGTGA	1381
Db	2424	TGGAACCTACTATTAACCAATGTATATCATCAATAAAGTCGTCCTTATAGTAGAGAGTGA	248
Qy	1381	AGAATTCTTAATGTGGAACTGCTGCCATGAAGAAAGACAGCCTATATGATGACCA	1441
Db	2484	AGAATTCTTAATGTGGAACTGCTGCCATGAAGAAAGACAGCCTATATGATGACCA	254
Qy	1441	CATGATGAAAACACTCTTGACTTATGGAATGGACGCGAAAGCCTATCTGCTTGCGCTCC	1501
Db	2544	CATGATGAAAACACTCTTGAGATTATGGAATGGACGCGAAAGCCTATCTGCTTGCGCTCC	260
Qy	1501	TCAGAGCTGTTAAACAGAGAACCTTACTATACAGACAGAGAAATTGAAACCCACATCAA	1561
Db	2604	TCAGAGCTGTTAAACAGAGAACCTTACTATACAGAGAGAGAAATTGAAACCCACATCAA	266
Qy	1561	GACCTCTCAATTTGTAGCACCTGATGAACATATTTGCTGCTATATGCGGTAAATATTCAT	1621
Db	2664	GACCTCTCAATTTGTAGCACCTGATGAACATATTTGCTGCTATATGCGGTAAATATTCAT	272
Qy	1621	GGCGTATGAGACAGGCTATTTCTAAACGCTGACACACACTGTGAGGAATGGGCTTACGCT	1681
Db	2724	GGCGTATGAGACAGGCTATTTCTTAACGCTGTCACACCACTGTGAGGAATGGGCTTACGCT	278
Qy	1681	CGCTGCCAAAGTTTACCGCTCTATGATGAGACCTACTGCTGTAAGAGAGCAATCCAGAA	1741
Db	2784	CGCTGCCAAAGTTTACCGGCTCTATGATGAGACCTACTGCTGTAAGAGAGCAATCCAGAA	284
Qy	1741	TTGGAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGGTCTCG	1801
Db	2844	TTGGAATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGGTCTCG	290
Qy	1801	TTTCAGGAGGAACCTCACCCTGCTCCACAAACCCCATCAACGAAAGTTTCAAGCTCAAC	1861
Db	2904	TTTCAGGAGGAACCTCACCCTGCTCCACAAACCCCATCAACGAAAGTTTCAAGCTCAAC	296
Qy	1861	ATCAGATAGTTCAACTTCACAGTCTAGCTTAACCACTCCAGACCAAAATATATGATGCAC	1921
Db	2964	ATCAGATAGTTCAACTTCACAGTCTAGCTTAACCACTCCAGACCAAAATATATGATGCAC	302
Qy	1921	TACCAATCTTAACAATTAATACGCAACATCAATTAACCCCTGATCAACAAATTCAGAA	1981
Db	3024	TACCAATCTTAACAATTAATACGCAACATCAATTAACCCCTGATCAACAAATTCAGAA	308
Qy	1981	TTCTCAACCAAGACACCA 1999	

Db	3084	TCCTCAACGACAGACCA	3102
RESULT 10	AE008414	10313 bp	DNA linear BCT 13-SEP-2001
LOCUS	AE008414/c		
DEFINITION	Streptococcus pneumoniae R6 section 30 of 184 of the complete genome.		
ACCESSION	AE008414	AE007317	
VERSION	AE008414.1	GI:15457885	
KEYWORDS			
SOURCE			
ORGANISM	Streptococcus pneumoniae R6. Streptococcus pneumoniae R6 Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.		
REFERENCE	1 (bases 1 to 10313) Hoskins,J.A., Alborn,W. Jr., Arnold,J., Blaszcak,L., Burgett,S., Dehoff,B.S., Estrem,S., Filiz,L., Fu,D.-J., Fuller,W., Geringer,C., Gilmour,R., Glass,J.S., Khoja,H., Kraft,A., Lagace,R., Leblanc,D.J., Lee,L.N., Leikowitz,E.J., Lu,J., Matsushima,P., Mcahren,S., McHenry,M., Mcleaster,K., Mundy,C., Nicas,T.I., Norris,F.H., O'Gara,M., Peery,R., Robertson,G.T., Rockey,P., Sun,P.-M., Winkler,M.E., Yang,Y., Young-Bellido,M., Zhao,G., Zook,C., Baltz,R.H., Jaskunas,S.R., Rostock,P.R. Jr., Skatrud,P.L. and Glass,J.I.		
AUTHORS	2 (bases 1 to 10313) Hoskins,J.A., Alborn,W. Jr., Arnold,J., Blaszcak,L., Burgett,S., Dehoff,B.S., Estrem,S., Filiz,L., Fu,D.-J., Fuller,W., Geringer,C., Gilmour,R., Glass,J.S., Hann,A., Khoja,H., Kraft,A., Lagace,R., Leblanc,D.J., Lee,L.N., Leikowitz,E.J., Lu,J., Matsushima,P., Mcahren,S., McHenry,M., Mcleaster,K., Mundy,C., Nicas,T.I., Norris,F.H., O'Gara,M., Peery,R., Robertson,G.T., Rockey,P., Sun,P.-M., Winkler,M.E., Yang,Y., Young-Bellido,M., Zhao,G., Zook,C., Baltz,R.H., Jaskunas,S.R., Rostock,P.R. Jr., Skatrud,P.L. and Glass,J.I.		
TITLE	Genome of the Bacterium Streptococcus pneumoniae Strain R6		
JOURNML	J. Bacteriol. 183 (19), 5709-5717 (2001)		
MEDLINE	21429245		
PUBMED	11544234		
FEATURES	1. 10313 /organism="Streptococcus pneumoniae R6" /strain="R6" /db_xref="taxon:171101" complement(114..2273) /gene="pdpA" /note="spr0329" complement(114..2273) /gene="pdpA" complement(114..2273) /gene="pdpA" complement(114..2273) /EC_number="2.3.2.-" /EC_number="2.4.1.129" /cdoon_start=1 /transl_table=1 /label=spr0329 /product="penicillin-binding protein 1A" /protein_id="AAK9133.1" /db_xref="GI:15457886" /translation="MNKPTLLRLIKYISLISFLVIAIYLGGVFFYYSKAPSLSE SKVATTSKITYDNKNQDLADISERVNQAANDIPDLKAVISIEDHFEFHRKID TIRILGAFRLNLSLOGSGALTQOLIKITLFSTSDQTSIRKAGQEWALAIQLEK ATROELITYINKYVMSGNGYMGQTAONAYGDLNLSLPOLALAGQAOVNOYD KSPDEAODRRLVYSEMKNQGYISAQYERKAYNPPTDLOSLSKASNPAYMDVL KEYINQVEETGYNLTTGMDVTTNDQEKHLMIDYNTDEYAPLDELGVEASTV DVANGKVIQAGARHSSNVSGINDAVENRDMGSMKPTIDYAPALEGYESTAT IHDPEPNYPGTNTPYVNMDBGYFGNTTQYALQGSNNPAAVEITLKNKYGNAKTEPLN GLGDEPSYHSNAISGNTESDQKRYGASSEKAAALAAAFANGVTYKPKYIKVVS DGESEKSNVGTFRAMKETTAYMMTDMMKTYLSYGTGNAYLAWLPQAGKGTGYNDE EIEHNKTSQFVAPDELFACTYTRYKISMAVWTGTSNRLTPLYGNGLYAAKYVSMTY LTSKSEPMEDNINIEGLYRNGEYFKNGARGATWSVPAPOQPTGSSSSSSSDSTSS		

gene  
STPSTNNSTTTNNPNNTTQOSNTTDDQONONPQAPQ"  
/complement(2270..2866)  
/gene="recu"  
/note="spr0330"  
/complement(2270..2866)  
CDS  
/gene="recu"  
/note="putative"  
/codon\_start=1  
/transl\_table=1  
/label-spr0330  
/product="Recombination protein u"  
/protein\_id="AAK9134.1"  
/db\_xref="GI:15457887"  
/translation="MVNYPKRVSSOKROTSLSOPKFNFRNGMSEKMINATNDYYSLO  
GLAVTHKKPTPIQIVQNDYVPORSRAKIYEAIFRQASTDYSGVINGYITIDEVKEIKO  
KRAIPMKNFHPHQIOHMEQVLQOQICVLLHFSSQOETYLIPAFDLIRFYHODGOK  
SMPELYIREYGEIKAGAFPOIPYLVNLIKELLGKTR"  
2932..3459  
CDS  
/gene="spr0331"  
2932..3459  
/gene="spr0331"  
/codon\_start=1  
/transl\_table=1  
/label-spr0331  
/product="Conserved hypothetical protein"  
/protein\_id="AAK9135.1"  
/db\_xref="GI:15457888"  
/translation="MATALVGYSAFDGLFSRDKPRLKIKKAIKRDLEAMADGVS  
WLVTGSLGEYVWLEVAQEMKTEYGFOLATIFAFETHGENMNGNOMKLSRFQKQVDF  
VKVAYPRHEKGOLRDYQOFLFLENTSSYLEFDEENETKLAVFYQKKMNQNDYFIKRL  
TFDOLNEAENFSEN"  
3517..3858  
gene  
/gene="spr0332"  
3517..3858  
/gene="spr0332"  
/codon\_start=1  
/transl\_table=1  
/label-spr0332  
/product="Conserved hypothetical protein"  
/protein\_id="AAK9136.1"  
/db\_xref="GI:15457889"  
/translation="MERDMSIIFRSANDIFPOEFGREYGVNKEVDFLDVYKDYE  
TYALVKSIRQELADLKEELRKPSPVQAEPLAATTSMTNFIDILKRLNLEKEY  
FGKQIIDNSD"  
3871..4274  
misc\_RNA  
/gene="rnpb"  
/product="Ribonuclease P - RNA component"  
3871..4274  
gene  
/gene="rnpb"  
4329..5501  
CDS  
/gene="spr0333"  
4329..5501  
/gene="spr0333"  
/codon\_start=1  
/transl\_table=1  
/label-spr0333  
/product="Conserved hypothetical protein"  
/protein\_id="AAK9137.1"  
/db\_xref="GI:15457890"  
/translation="MODCNMKKEFNLIATYVAGLEAVYGVRELYGDCOVENGVRP  
QGVRAIITENMLARADRIKIITGTFPAKTFEELFGCVFALDMENTLPLDARFPIK  
AKCVKSLHNEPVSQALSKAVNKKLOKHARPGVPLMENGPEFKLEVSILKQVATY  
MIDITGSSLEKRGYRTERKGAPIKENMAAAILQLSNNYPDKPLIDPTGSGTFCIEAV  
MIRKMAAGLGRSFAPFEMWMSIDRLIQEVETEAKEYERLELDIMGCDIDARVEI  
AKANAQAGVAGDITFCOMRVODLRSDKINVIISNPPYGERLSDDAGVTKLYAEMGO  
VEAPLKTMSKFIILSDAEFSKYSQADKKKKLNLGTLKVLVLYYFGQRYKQEVK"  
5514..6908  
gene  
/gene="spr0334"  
5514..6908  
/gene="spr0334"  
/codon\_start=1  
/transl\_table=1  
/label-spr0334

/product="Conserved hypothetical protein"  
/protein\_id="AAK9138.1"  
/db\_xref="GI:15457891"  
/translation="MSKRRNRHKKEAQEPQDFDEAKELTVGAIRKNEVEAGVLP  
EDSILDKYKQKHDEIADKFAIRQYKEEVEIQSLDLDLQEMEAFAVEKSEASEEV  
PSSBIDILPLDDEEGDLPDLDDENPTMETVEEQQNLSRQDSEKSKGKGF  
ILVLAIVSVIICVSAVYVROVARSTKEIETSSQTANQSDVDENFLYAPFTNSN  
KTKALNSQFDKLSQKTLTKLDEGRRETTAKSYDSLATQIKAOVNAQPEKPAIV  
DGVLDITNAKAKSDAKFTDIKTGNTELDKVLKALISLGSQQTSTSSSSSQTSSSSS  
QASSNTSEKPPSSNETRSRSREVMGLSAGAVAVQASRAVAIQAOSADIDSSNSAW  
DFADGVLEQIILATSKSRGCIIGDQYILERVNIIVNGNYYNLKPKGYTLFLNCKTGY  
FVGNAGHADDLDY"  
6984..8408  
gene  
/gene="gnd"  
/note="spr0335"  
6984..8408  
CDS  
/gene="gnd"  
/EC\_number="1.1.1.44"  
/codon\_start=1  
/transl\_table=1  
/label-spr0335  
/product="6-phosphogluconate dehydrogenase"  
/protein\_id="AAK9139.1"  
/db\_xref="GI:15457892"  
/translation="WTKANPEYGVGAWVRMLAINIESRGYTVAILNRSKRETDVIA  
CHPEKNFVSTDESFEVNSIEKPRIMLMVQAGPDTFTQALLPHLDKGLDIDGN  
TFYKDTIRRNBELANSQINFIGTGVSGEGALGSPSLPESGOEAYELVADVLEIS  
AKADEDKPCVYTIIGPDAGHYKVMHNGIIEYGMQLIASEYDLMQHLGLSAEEMAR  
IETFMNKELSYLIETIADILSRKDDGGDGPVYDYLDAAGNKGTEKWTQSQSLD  
GVLSTLSEVFARYIETYEKREYHASKVLPKPAAPNEGSKATEIKETIKATPESKI  
ISYAGFPALQVASKENNMNLPFADIASIMPDCGCIIRSRPQKLTDAIENRADLANLL  
IDETFLDYTAQYQQAVRDVVALAVQAGPVTFSAITYFDSIRASADLPANLQARD  
YFGAHTYQKDKDEGTFHYSWYDEK"  
8420..9109  
gene  
/gene="csr"  
8420..9109  
CDS  
/gene="csr"  
/note="putative"  
/codon\_start=1  
/transl\_table=1  
/label-spr0336  
/product="response regulator"  
/protein\_id="AAK9140.1"  
/db\_xref="GI:15457893"  
/translation="WGRKRIILKEERNLAHFLSLQKEQYRVDLVEGOKALSMAIQ  
TDVYLILNVNLGDMQDFAEKLSRTKPAVIMILDHMEDLOELEEVORFAVSITY  
KPVLIENVARISALFRGRDPTIDHCSIMKAPRYRNIRIDVEHHYVRGEMATLR  
REYDLATLMSKRVLTREQLLESVWKIESATETNIYDVYIRILRSKLIDVAGOKSYIK  
TVRGVGYTMOE"  
9207..10223  
gene  
/gene="cbpf"  
/note="spr0337"  
9207..10223  
CDS  
/gene="cbpf"  
/codon\_start=1  
/transl\_table=1  
/label-spr0337  
/product="Choline-binding protein F"  
/protein\_id="AAK9141.1"

Query Match 99.1%; Score 1981.4; DB 1; Length 10313;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTACGCAATAAATCACTGCTGACTTGGTTCGTGAACGCCGGCTCAA 60  
|||||  
DB 2115 TAAATCTACGCAATAAATCACTGCTGACTTGGTTCGTGAACGCCGGCTCAA 2056  
|||||  
QY 61 TGCCCAAGCTAATGATATTCACAGATTTGGTAAAGCAATCTTCTATGAGAACCA 120  
|||||  
DB 2055 TGCCCAAGCTAATGATATTCACAGATTTGGTAAAGCAATCTTCTATGAGAACCA 1996  
|||||

QY 121 TCGCTTCTGACACACAGGGGATGATACATCCGTAATCTGGAGCTTCTTGCCGAA 180  
 |||||||  
 Db 1995 TCGCTTCTGACACACAGGGGATGATACATCCGTAATCTGGAGCTTCTTGCCGAA 1936  
 |||||||  
 QY 181 TCTGAAAGCAATTCCTCCAGAGGATGATCACTCCACCAAGCTGATTAAGTTAG 240  
 |||||||  
 Db 1995 TCTGAAAGCAATTCCTCCAGAGGATGATCACTCCACCAAGCTGATTAAGTTAG 1876  
 |||||||  
 QY 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGAAAGCTTGT 300  
 |||||||  
 Db 1875 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTGTAAGGCTCAGAAAGCTTGT 1816  
 |||||||  
 QY 301 AGCGATTAGTTAGAACAAAAAGACACACAGAAATCTTGACTCTCTATTAATTA 360  
 |||||||  
 Db 1815 AGCGATTAGTTAGAACAAAAAGACACACAGAAATCTTGACTCTCTATTAATTA 1756  
 |||||||  
 QY 361 GGTCTACATGCTAATGAGGACTATGAAATGACAGACAGCTCAAACTATAGTAA 420  
 |||||||  
 Db 1755 GGTCTACATGCTAATGAGGACTATGAAATGACAGACAGCTCAAACTATAGTAA 1696  
 |||||||  
 QY 421 AGACCTCAATTAATTAAGTTTACCTCACTAGCTTGGCTGGATGACCTCAGGAC 480  
 |||||||  
 Db 1695 AGACCTCAATTAATTAAGTTTACCTCACTAGCTTGGCTGGATGACCTCAGGAC 1636  
 |||||||  
 QY 481 AAACCAATATGACCCCTATTACATCCAGAGACGCCAAGCCGGAATGCTGTT 540  
 |||||||  
 Db 1635 AAACCAATATGACCCCTATTACATCCAGAGACGCCAAGCCGGAATGCTGTT 1576  
 |||||||  
 QY 541 ATCTGAATGAAAAATCAAGGCTACATCTGCTGAAACAGTATGAAAGAGCAATAC 600  
 |||||||  
 Db 1575 ATCTGAATGAAAAATCAAGGCTACATCTGCTGAAACAGTATGAAAGAGCAATAC 1516  
 |||||||  
 QY 601 ACCAATTAATGATGAGTCAAAAGTCTCAATTCAGCAAGTATTAACCTGCTTACATGA 660  
 |||||||  
 Db 1515 ACCAATTAATGATGAGTCAAAAGTCTCAATTCAGCAAGTATTAACCTGCTTACATGA 1456  
 |||||||  
 QY 661 TAAATTAATGATGAGTCAAAAGTCTCAATTCAGCAAGTATTAACCTGCTTAC 720  
 |||||||  
 Db 1455 TAAATTAATGATGAGTCAAAAGTCTCAATTCAGCAAGTATTAACCTGCTTAC 1396  
 |||||||  
 QY 721 AACTGGATGATGCTCAAAATGTAGACCAAGAACTCAAAACCTGCTGGATAT 780  
 |||||||  
 Db 1395 AACTGGATGATGCTCAAAATGTAGACCAAGAACTCAAAACCTGCTGGATAT 1336  
 |||||||  
 QY 781 TTACAATACAGACGATACGTTGCTATCCAGAGATGAATGGCAAGTCTTACCAT 840  
 |||||||  
 Db 1335 TTACAATACAGACGATACGTTGCTATCCAGAGATGAATGGCAAGTCTTACCAT 1276  
 |||||||  
 QY 841 TGTGATGATTTCTAAGCTTAAGTCAATGGCCAGCTAGGACGCCATCAGTCAAGTAA 900  
 |||||||  
 Db 1275 TGTGATGATTTCTAAGCTTAAGTCAATGGCCAGCTAGGACGCCATCAGTCAAGTAA 1216  
 |||||||  
 QY 901 TGTTCCTTGGGAATTAACGAGATGAAACAAACCGGAGCTGGGATCAACATGAA 960  
 |||||||  
 Db 1215 TGTTCCTTGGGAATTAACGAGATGAAACAAACCGGAGCTGGGATCAACATGAA 1156  
 |||||||  
 QY 961 ACCGATACAGACTATGCTTGCCTTGAGTAGGCTTACAGTTCATGCTACTAT 1020  
 |||||||  
 Db 1155 ACCGATACAGACTATGCTTGCCTTGAGTAGGCTTACAGTTCATGCTACTAT 1096  
 |||||||  
 QY 1021 CGTTCAGATGAGCCCTTAATCACTACCTGGGACAAATCTCTGTTTATTAAGGAGTAG 1080  
 |||||||  
 Db 1095 CGTTCAGATGAGCCCTTAATCACTACCTGGGACAAATCTCTGTTTATTAAGGAGTAG 1036  
 |||||||  
 QY 1081 GGGCTACTTGGCAACATACCTTGCAATAGCCCTGCAAAATCGGCAACGTCGCCAGC 1140  
 |||||||  
 Db 1035 GGGCTACTTGGCAACATACCTTGCAATAGCCCTGCAAAATCGGCAACGTCGCCAGC 976  
 |||||||  
 QY 1141 CGTGAAGACTTAACAAGTTCGACTCAACGGGCAAGCTTCTTAATGCTGTAAG 1200  
 |||||||  
 Db 975 CGTGAAGACTTAACAAGTTCGACTCAACGGGCAAGCTTCTTAATGCTGTAAG 916  
 |||||||  
 QY 1201 AATGACTACCAAGATTAATCACTACTCAATGCCATTTCAAGTAACACACGGAATCAGA 1260  
 |||||||

Db 915 AATGACTACCAAGATTAATCACTACTCAAAATGCCATTTCAAGTAACACACGGAATCAGA 856  
 |||||||  
 QY 1261 CAAAAATATGAGACAAAGTATGAAAAAGATGGCTGCTTACGCTGGCTTTGCAAAATGG 1320  
 |||||||  
 Db 855 CAAAAATATGAGACAAAGTATGAAAAAGATGGCTGCTTACGCTGGCTTTGCAAAATGG 796  
 |||||||  
 QY 1321 TGAACCTTAATTAACCAATGATATTCATTAAGTCTGTTTATGATGAGGAGTGA 1380  
 |||||||  
 Db 795 TGAACCTTAATTAACCAATGATATTCATTAAGTCTGTTTATGATGAGGAGTGA 736  
 |||||||  
 QY 1381 AGAGTTCTTAATGTCGGAACCTGCGCATATGAGAAACGACGCTTATGATGACCGA 1440  
 |||||||  
 Db 735 AGAGTTCTTAATGTCGGAACCTGCGCATATGAGAAACGACGCTTATGATGACCGA 676  
 |||||||  
 QY 1441 CATGATGAAGACAGCTTGAATTTATGAACTGAGAGAAATCCCTATCTTGGCTGCC 1500  
 |||||||  
 Db 675 CATGATGAAGACAGCTTGAATTTATGAACTGAGAGAAATCCCTATCTTGGCTGCC 616  
 |||||||  
 QY 1501 TCAGGCTGTAAACAGGAACTCTAATATACAGACGAGAAATGGAACACATCA 1560  
 |||||||  
 Db 615 TCAGGCTGTAAACAGGAACTCTAATATACAGACGAGAAATGGAACACATCA 556  
 |||||||  
 QY 1561 GACCTCTAATTTAGTACGCTGATGAACATTTTGTGCTATATACGCTTAATATTCAT 1620  
 |||||||  
 Db 555 GACCTCTAATTTAGTACGCTGATGAACATTTTGTGCTATATACGCTTAATATTCAT 496  
 |||||||  
 QY 1621 GCGTGTATGACAGGCTTTCTAACCCTGTCACACCACTTATAGCAATGGCCTTACG 1680  
 |||||||  
 Db 495 GCGTGTATGACAGGCTTTCTAACCCTGTCACACCACTTATAGCAATGGCCTTACG 436  
 |||||||  
 QY 1681 CGTGCCAAGTTTACCGCTTATGATGACCTTACCTGCTGTAAGGAACATCCAGAA 1740  
 |||||||  
 Db 435 CGTGCCAAGTTTACCGCTTATGATGACCTTACCTGCTGTAAGGAACATCCAGAA 376  
 |||||||  
 QY 1741 TTGAATTAACCAAGGCTCTACAGAAATGGAATTCCTATTTAAATGGTGTGCG 1800  
 |||||||  
 Db 375 TTGAATTAACCAAGGCTCTACAGAAATGGAATTCCTATTTAAATGGTGTGCG 316  
 |||||||  
 QY 1801 TTACGATGAACTACCTGCTCCACAACACCCCATCACTGAAAGTTCAACCTATC 1860  
 |||||||  
 Db 315 TTACGATGAACTACCTGCTCCACAACACCCCATCACTGAAAGTTCAACCTATC 256  
 |||||||  
 QY 1861 ATCAGATAGTTCAACTTACAGTGTAGCTCAACCACTCCAGCAAAATATATAGTAC 1920  
 |||||||  
 Db 255 ATCAGATAGTTCAACTTACAGTGTAGCTCAACCACTCCAGCAAAATATATAGTAC 196  
 |||||||  
 QY 1921 TACCAATCTTAACAATATATACGCAACAATCAATACCAACCTGATCAACAATTCAG 1980  
 |||||||  
 Db 195 TACCAATCTTAACAATATATATACGCAACAATCAATACCAACCTGATCAACAATTCAG 136  
 |||||||  
 QY 1981 TCTCAACGACGACACCA 1999  
 |||||||  
 Db 135 TCTCAACGACGACACCA 117  
 |||||||

RESULT 11  
 AF210752 2160 bp DNA linear BCT 17-SEP-2001  
 LOCUS  
 DEFINITION Streptococcus pneumoniae isolate BM 4200 penicillin-binding protein  
 1A (pbp1a) gene, complete cds.  
 ACCESSION AF210752  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Streptococcus pneumoniae.  
 Streptococcus pneumoniae  
 Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 Streptococcus.  
 REFERENCE  
 1 (bases 1 to 2160)  
 Authors Perroni, A. and Berche, P.  
 Title Alterations to penicillin-binding proteins 1A, 2B and 2X amongst  
 penicillin-resistant clinical isolates of Streptococcus pneumoniae  
 serotype 23f from the nasopharyngeal flora of children



Db	1659	ACAACGAGTAAACAGGACTTCTTAACGATATCTGACGAGAAATTGAAAGCTATATCA	1718
QY	1561	GACCTCTCAATTTGTAGCACCTGATGAACATATTGCTGGCTATACGCGTAATATTCAAT	1620
Db	1719	GAAACCTGGTTACGTAGCTCCAGATGAATGTTTGTAGGGTATPACCCGCAATATGCCAAT	1778
QY	1621	GCCCTATGAGACAGGCTATPCTAACCGTGTGACACACTGTTGAGGCAATGGCGTTACGGT	1680
Db	1779	GGCTTATGAGACAGGCTATTCTTAACCGTGTGACACCACTGTGAGCGATGAGGCGCTTACGGT	1838
QY	1681	CGCTCCCAAAGTTTACCGCTCTATGATGACCCACTGCTGTGGAAGGAAGCAATCCAGAGA	1740
Db	1839	CGCTCTAAAGTTTACCGCTCTATGATGACCCACTGCTGTGGAAGGAAGCAATCCAGAGA	1898
QY	1741	TTGGAATATTACCAAGGGGCTCTACAGAAATGAGCAATTCGATTTTAAAAATGCTGCTCG	1800
Db	1899	CTGGAAATATTACCAAGGGGCTCTACAGAAATGAGCAATTCGATTTTAAAAATGCTGCTCG	1958
QY	1801	TTTCTACGGGAACTACCGTGCCTCCACAACACCCCAATCAACTGAAAGTCAAGCTCATTC	1860
Db	1959	TTTCTACGGGAACTACCGTGCCTCCACAACACCCCAATCAACTGAAAGTCAAGCTCATTC	2018
QY	1861	ATCAGATAGTTCACATCTCACAGTCTAGCTCAACCACTCCAAAGCACAATAATATGATGACAC	1920
Db	2019	ATCAGATAGTTCACATCTCACAGTCTCACTCAACCACTCCAAAGCACAATAATATGATGACAC	2078
QY	1921	TACCAATCCTAACAAATTAATAGCAACAATCAAAATACAACCCCTGATCAACAAATTCAGAA	1980
Db	2079	TACCAATCCTAACAAATTAATAGCAACAATCAAAATACAACCCCTGATCAACAAATTCAGAA	2138
QY	1981	TCCTTAACCCAGACACCA 1999	
Db	2139	TCCTTAACCCAGACACCA 2157	

RESULT	12
AF139884	
LOCUS	2160 bp DNA linear BCT 26-AUG-1999
DEFINITION	Streptococcus pneumoniae strain PO-273 penicillin-binding protein
ACCESSION	1a (pbpla) gene, complete cds.
VERSION	AF139884
KEYWORDS	AF139884.1 GI:5410458
SOURCE	
ORGANISM	Streptococcus pneumoniae. Streptococcus pneumoniae Bacteria; Filimicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.
REFERENCE	1 (bases 1 to 2160)
AUTHORS	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
TITLE	Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of Streptococcus pneumoniae arose by large recombinational replacements of the cpsA-pbpIIa region Microbiology 145 (Pt 8), 2023-2031 (1999)
JOURNAL	
MEDLINE	99392464
PUBMED	10463168
REFERENCE	2 (bases 1 to 2160)
AUTHORS	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
TITLE	Direct Submission
JOURNAL	Submitted (31-MAR-1999) The Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford, Oxfordshire OX1 3PS, UK
FEATURES	
source	location/Qualifiers 1..2160 /organism="Streptococcus pneumoniae" /strain="PO-273" /serotype="14" /db_xref="taxon:1313" /note="Polish variant of the major penicillin-resistant Spanish serotype 9V clone" 1..2160 /gene="pbpla"
gene	

```

CDS
1..2160
/gene="hnp1a"
/function="peptidoglycan transglycosylase-transpeptidase"
/codon_start=1
/trans_table=11
/product="penicillin-binding protein 1a"
/protein_id="A043067.1"
/db_xref="GI:5410459"
/translation="MKPRTLRILKYLISFSLVIAVAIVLGGVVEFYVSKAPSLSE
SKIVATTSKRIYDKRNKLIADIAGERVRNQAQNDIPDLKAVISIEDHFAEPHGRGD
SIRLIGAFESKRIYDLOSNGSGSTLQOCLIKLTVEYSPDOTIRKQAEALVALOEOK
ATRECLITLYINKYVMSNGNYGMQTAQNYANGDIINLSTPOLAILAGMPOARNOAYL
KSPHEAQQDRNRNLYLSMKNGCYTSAQCYKAYKATPRTSLDLSLKSANPRAQMDYL
KEYNOYEOETGILTLTGMDVYTNVDEQKHLMIDYINSDQYVSFDDDLQYASTYV
DVNSGKVIADQARHQASNVSFQNCQVEENRQMSAKRTIDYAPRIETGVITDSTL
MVNDIPYNPQSTIPRYNMDRAVFGNTTIDALQOCSNVAETLKNVGLGDRKTLN
GLGIDYSPMHYANALISNTTESNKOYASSEKMAAALAAANGTYKKPYIKRVYS
DGSKEESNVGTRAMKEETAYVMNTDMKATVLVTGCTGAYLPWLPOAGKGTGNTYDE
EVEEHKIKGTGVAPEDEMFGVETPKRYSMAVVTGSNRLTPIVGDGFLAAKAVYRSMY
LSSESNPEDMINPEGLIYRNGEYFAPKNASTWSPADQOOPRESSSSSSSDSTSQSS
STPSTSNSTTTPNNNTQOSNTTPDQONQNPQAP"
BASE COUNT
693 a      494 c      407 g      566 t
ORIGIN

```

Query Match	85.3%	Score 1704.6	DB 1	length 2160	
Best Local Similarity	90.8%	Pred. No. 0			
Matches 1815	Conservative	0	Mismatches 184	Indels 0	Gaps 0

QY	1	TAAATCTACACAAATAAATAACAACTCATCTGCTGACCTTGGGTTCTGTAACGCCCGTCAA	60
DB	159	TAAATCTACACAAATAAATAACAACTCATCTGCTGACCTTGGGTTCTGTAACGCCCGTCAA	218
QY	61	TGCCCAAGCTAATGATATTTCCACAGATTGGTTAAGGCAATCGTTTCTATGAAAGCA	120
DB	219	TGCCCAAGCTAATGATATTTCCACAGATTGGTTAAGGCAATCGTTTCTATGAAAGCA	278
QY	121	TGCGTCTTCGACACCAAGGGGGATGATACCATCCGTAATCTGGGAGGTTCTTGCCCAA	180
DB	279	TGCGTCTTCGACACCAAGGGGGATGATTCATCCGTAATCTGGGAGGTTCTTGCCCAA	338
QY	181	TCTGCAAGCAATTCGCCCAAGGTGATCTCAACTCTCCACCAAGTTGATTAAGTTGAC	240
DB	339	TCTGCAAGTAAATTCGCCCAAGGTGATCTCAACTCTCCACCAAGTTGATTAAGTTGAC	398
QY	241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTTCTGTAAGCTGAGAAAGCTTGGTT	300
DB	399	TTACTTTTCAACTTCGACTTCGACACAGACTATTTTCTGTAAGCTGAGAAAGCTTGGTT	458
QY	301	AGCGATTGACGTTAGAACAAAAAAGAACCAACCAAGCAAGAAATCTTGACACTATATAATTA	360
DB	459	AGCGATTGACGTTAGAACAAAAAAGAACCAACCAAGCAAGAAATCTTGACACTATATAATTA	518
QY	351	GGTCACACTGCTAATGGGAAGCTATGGAAATGCACAGACCTCAAAAACCTAATGGTAA	420
DB	519	GGTCACACTGCTAATGGGAAGCTATGGAAATGCACAGACCTCAAAAACCTAATGGTAA	578
QY	421	AGACCTCAATATTTAATGATTACCTCAAGTTAGCCTTGCTGGCTGGAATGCTCAGGCACC	480
DB	579	AGACCTCAATATTTAATGATTACCTCAAGTTAGCCTTGCTGGCTGGAATGCTCAGGCACC	638
QY	481	AAACCAATATGACCCCTAATTCACATCCAGAAAGACGCCAAGCCGCGAAACCTTGCTTT	540
DB	639	AAACCAATATGACCCCTAATTCACATCCAGAAAGACGCCAAGCCGCGAAACCTTGCTTT	698
QY	541	ATCTGAAATGAAAAATCAGGCTACATCTCTGCTGAAACAGTATGAGAAAGCAGTCAATAC	600
DB	699	ATCTGAAATGAAAAATCAGGCTACATCTCTGCTGAAACAGTATGAGAAAGCAGTCAATAC	758
QY	601	ACCAATTAAGTATGAGCTACAAAGTCTCAAAATGAGCAAGTAATTAACCTGGTTACATGGA	660
DB	759	ACCAATTAAGTATGAGCTACAAAGTCTCAAAATGAGCAAGTAATTAACCTGGTTACATGGA	818

661 TAATTACCTCAAGAGATCATCAATCAAGTTGAAAGAAAGAGCGCTATTAACCTACTCAC 720  
|||||  
Db 819 TAATTACCTCAAGAGATCATCAATCAAGTTGAAAGAAAGAGCGCTATTAACCTACTCAC 878  
721 AACTGGAGATGATGCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780  
|||||  
Db 879 TACTGGAGATGATGCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 938  
781 TTACATATACAGACGAATAGCTGCTATCCAGACATGAATTGCAAGTGCCTTCTACCAT 840  
|||||  
Db 939 CTACACAGCTCCGATCAATAGCTCTTACCTGACGATGATTTGCAAGTGCCTTCTACCAT 998  
941 TGTATGATGTTTCAACGGTAAAGATGATCCAGCTAGAGAGACCGCATCACTCAAGTAA 900  
|||||  
Db 999 CGTACATGTTTCAAAATGGTAAAGATGATCCAGCTAGAGAGACCGCATCACTCAAGTAA 1058  
901 TGTTCCTTGGGAATTAACCAAGAGTAAAGCAAAACCGGATGGGATCACTATGAA 960  
|||||  
Db 1059 CGTTTCATTTGGTACCAACCAAGCTGTGGAACCAATCGTACCTGGGATTCGTATGAA 1118  
961 ACCGATTCACAGATCATGCTCCCTGCTGAGTACGCTGCTACGATTCACCTGCTACTAT 1020  
|||||  
Db 1119 ACCAATTCACCGATTAATGACCTGCTGATGAAATGCTTTATGATTCACCTGCTACTAT 1178  
1021 CGTTACGATGAGCCCTATTAATCACTACCTGGACAAATATCTGTTTATTAATGGGATAG 1080  
|||||  
Db 1179 GGTATGATGATATCTCTTATTAATCACTGCGGAAACCAAGACACCTGCTACACATGGGATAG 1238  
1081 GGGTACTTTGGCAATCAACCTGCTGCAATACGCTGCAACAAATGCGAAGCGTCCAC 1140  
|||||  
Db 1239 AGCATATTTGGTATTAATTAATCTGCTGCAATATGCTCTTCACAAATGCAAGAAATGTCACAC 1298  
1141 CGTGAACATCTAAACAAAGTGGGATCAACCGCCCAAGACTTCTTAAATGGTCTAG 1200  
|||||  
Db 1299 CGTTGACATTTGGAATGATAGTGGTCTGATAGTACCTAAACCTTCTTAAATGGTCTAG 1358  
1201 AATGACATACCAAGATTAATCACTACCTGCAATGCAATTCCTTAACTAACCAAGGATTCGA 1260  
|||||  
Db 1359 TATGACATATCCAAATGATGATTAATGCAACCGCATTTTCAATTAATACACAGAAATCTAA 1418  
1261 CAAAAATATGAGAGCAAGTATGTAAGAAAGTGGTCTGCTTACGCTGCTTTCGCAATAG 1320  
|||||  
Db 1419 TAAACATATGAGAGCAAGTATGTAAGAAAGTGGTCTGCTTACGCTGCTTTCGCAATAG 1478  
1321 TGGAACTTACTTAACCAATGATATATCATTAAGTCTGCTTATGATGAGAGTAA 1380  
|||||  
Db 1479 TGGCACTTACTTAACCAATGATATATCATTAAGTCTGCTTACGCTGATGAGTAA 1538  
1381 AGAGTTCTCTAATGTCGGAACCTGTCGCTGTAAGGAAAGACAGGCTATATGATGACGA 1440  
|||||  
Db 1539 AGAGTTCTCTAATGTCGGAACCTGTCGCTGTAAGGAAAGACAGGCTATATGATGACGA 1598  
1441 CATGATGAAACAGTCTTGAATTAAGAACTGAGCAAAATGCTTATGCTTGGCTGCC 1500  
|||||  
Db 1599 CATGATGAAACAGTCTTGAATTAAGAACTGAGCAAAATGCTTATGCTTGGCTGCC 1658  
1501 TCAGAGCTGTTAAACAGGACCTCTAATCTATACAGACAGGAAATTTGAAACACATCAA 1560  
|||||  
Db 1659 TCAGAGCTGTTAAACAGGACCTCTAATCTATACAGATGAGGAAATTTGAAACACATCAA 1718  
1561 GACCTCTCAATTTGAGCACTGATGAACTATTTGCTGCTATACGCTTAATATTTCAAT 1620  
|||||  
Db 1719 GACCACTGCTATGATGCTGCAATGATGAAATGTTTGTGTTATCTGTTATCTGTTATCTAT 1778  
1621 GCGTGTATGAGAGGCTATTTCAACGCTGTCGACCACTTGTAGGCAATGCGCTTACGCT 1680  
|||||  
Db 1779 GCGTGTATGAGAGGCTATTTCAACGCTGTCGACCACTTGTAGGCAATGCGCTTACGCT 1838  
1681 CCGTCCCAAGTTTACCGCTCTATGATGACCTACTGCTCTGAAGAGCAATTCAGAGA 1740  
|||||  
Db 1839 TGCACCTAAAGTTTATCGCTCTATGATGACCTACTGCTCTGAAGAGCAATTCAGAGA 1898  
1741 TTGGAATATACAGAGGCGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGGTGCCTG 1800

|||||  
Db 1899 TTGGAATATACAGAGGCGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGGTGCCTG 1958  
1801 TTCTACGTGGAATCACTGCTCTCACAAACACCCCATCACTGAAAGTTCAAGCTCATC 1860  
|||||  
Db 1959 TTCTACGTGGAATCACTGCTCTCACAAACACCCCATCACTGAAAGTTCAAGCTCATC 2018  
1861 ATCAGATGTTCACTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
|||||  
Db 2019 ATCAGATGTTCACTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2078  
1921 TACCAATCTCAATTAATATACCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1980  
|||||  
Db 2079 TTACCAATCTCAATTAATATACCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2138  
1981 TCCTCAACCGACCAACCA 1999  
|||||  
Db 2139 TCCTCAACCGACCAACCA 2157  
RESULT 13  
AF139885 2160 bp DNA linear BCT 26-AUG-1999  
LOCUS  
DEFINITION  
1a (pbp1a) gene, complete cds.  
ACCESSION  
AF139885  
VERSION  
AF139885.1 GI:5410460  
KEYWORDS  
SOURCE  
ORGANISM  
Streptococcus pneumoniae.  
Streptococcus pneumoniae.  
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
Streptococcus.  
1 (bases 1 to 2160)  
REFERENCE  
AUTHORS  
Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.  
TITLE  
Sero-type 14 variants of the Spanish penicillin-resistant serotype  
9V clone of Streptococcus pneumoniae arose by large recombinational  
replacements of the cpsA-pbp1a region  
JOURNAL  
Microbiology 145 (Pt 8), 2023-2031 (1999)  
99392464  
MEDLINE  
10463168  
PUBMED  
2 (bases 1 to 2160)  
REFERENCE  
AUTHORS  
Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (31-MAR-1999) The Wellcome Trust Centre for the  
Epidemiology of Infectious Disease, Department of Zoology,  
University of Oxford, South Parks Road, Oxford, Oxfordshire OX1  
3PS, UK  
FEATURES  
source  
Location/Qualifiers  
1. 2160  
/organism="Streptococcus pneumoniae"  
/strain="po-341"  
/serotype="14"  
/db\_xref="taxon:1313"  
/note="Polish variant of the major penicillin-resistant  
Spanish serotype 9V clone"  
1. 2160  
/gene="pbp1a"  
1. 2160  
/gene="pbp1a"  
/function="peptidoglycan transglycosylase-transpeptidase"  
/codon\_start=1  
/transl\_table=11  
/product="penicillin-binding protein 1a"  
/protein\_id="AADA3068.1"  
/db\_xref="GI:5410461"  
/translation="MNKPTILRLKYSISFLSVIAIVLGGVFYYSKAFSLSE  
SKIIVATSSKILYDNKNLIDIGSRVRNAQNDIPDVLKVAIYSIEDHRRPDRGID  
SIRLIGATRLRLQNSLOGSSTLRDQILKLYFSTSTSDQTSKRAQDAWLAIOLEOK  
ATKQDILITTYITNKYVMSNGNEMQTAQNTYTGKIDNLSLFDLALLAMPOAPMDIDP  
YSHPAADRRNLVISEKNGYISAEYEAQVNPITDGLQSLKSASNPAYMDNYL  
KVINVOEETGYMLTTPGMVYNNVDOEAKHLMIDYNSQOYVSPDDDLQVASTVY  
DVNSGKVIAGQARHQASNVSEFTNQAQVETRWDSAAKPTDYAPALEYGVYSTAT  
MVNDIPYNYPETSTPYVNMWRAYGNITLQYVALQDSNRVTAVERLTKVGLDRATFLN



GLGIDYPSMHYANAIISSNTTESNKQYASSSKMAAAYAFNGSTYYIKPMYIHKVES  
DGSKKERSNVGTRAMKETETATAMTDIMKTALVLTJYTGRAAY.PMLPQAGKGTSTYDE  
EVENHINNTGYAPDEMFVGTATRYKSYAENVNRLPTIVGDEFLLAAKAYRSMYTY  
LSEGSNEDMANIPEGLYRNGEFVKNSGRSFTWSSPADPOPSTESSSSSSSSSSSS  
STPSTJNNSTTNNNNNTQOSNTTPODONONPOQAQ<sup>8</sup>

BASE COUNT 693 a 494 c 407 g 566 t

ORIGIN

Query Match 85.3%; Score 1704.6; DB 1; Length 2160;  
Best Local Similarity 90.8%; Pred. No. 0;  
Matches 1815; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

QY 1 TAAAAATACGACAAATYAAAAATCAATCATCTGCTGACCTGGGTTCTGAAGCCGCGTCAA 60  
Db 159 TAAAAATACGACAAATYAAAAATCAATCATCTGCTGACCTGGGTTCTGAAGCCGCGTCAA 218  
QY 61 TGCCCAACCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120  
Db 219 TGCCCAACCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278  
QY 121 TCGCTTCTTCGACACAGAGGGGATGATACCATCCGTATCCGGAGGCTTCTTCGCGAA 180  
Db 279 TCGCTTCTTCGACACAGAGGGGATGATACCATCCGTATCCGGAGGCTTCTTCGCGAA 338  
QY 181 TCTGCAAGCAATTCCTCCAGAGTGATCAACTCCACCAACAGTTGATTAACTGAC 240  
Db 339 TCTGCAAGCAATTCCTCCAGAGTGATCAACTCCACCAACAGTTGATTAACTGAC 398  
QY 241 TTACTTTCAACTTGACATCTCGACACGATTTTCTGTAAGGCTCGAAGAGCTTGTT 300  
Db 399 TTACTTTCAACTTGACATCTCGACACGATTTTCTGTAAGGCTCGAAGAGCTTGTT 458  
QY 301 AGCGATTAGTTAGAACAAAAAGCAACAGCAAGAAATCTTGACCTATATATATATA 360  
Db 459 AGCGATTAGTTAGAACAAAAAGCAACAGCAAGAAATCTTGACCTATATATATATA 518  
QY 361 GGTCTACATGCTAATGGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 420  
Db 519 GGTCTACATGCTAATGGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 578  
QY 421 AGACCTCAATATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 480  
Db 579 AGACCTCAATATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 638  
QY 481 AAACCAATATGACCCCTATTCACATCCAGAAAGAGCCCAAGAGCCGCAAACTTGCTT 540  
Db 639 AAACCAATATGACCCCTATTCACATCCAGAAAGAGCCCAAGAGCCGCAAACTTGCTT 698  
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTGCTGAAACAGTATGAGAAAGAGTCAATAC 600  
Db 699 ATCTGAAATGAAAAATCAAGGCTACATCTGCTGAAACAGTATGAGAAAGAGTCAATAC 758  
QY 601 ACCAATTTCTATGACATCAACAAAGTCTCAATTCAGCAAGTATTTACCTGCTTACATGGA 660  
Db 759 ACCAATTTCTATGACATCAACAAAGTCTCAATTCAGCAAGTATTTACCTGCTTACATGGA 818  
QY 661 TAAATTAACCTCAAGAGAGTCAATCAAGTGAAGAGAAAGAGAGGCTTAATCACTACTCAC 720  
Db 819 TAAATTAACCTCAAGAGAGTCAATCAAGTGAAGAGAAAGAGAGGCTTAATCACTACTCAC 878  
QY 721 AACTGGGATGATGTTCTACACAAATGTAGACCAAGAGCTCAAAAAACATCTGTGGGATAT 780  
Db 879 TACTGGGATGATGTTCTACACAAATGTAGACCAAGAGCTCAAAAAACATCTGTGGGATAT 938  
QY 781 TTACAATACACACGAATACGTTGCTATCCAGAGATGAATTTGCAAGTCGCTTACATAT 840  
Db 939 CTACAACCTCGATCAATAGCTCTCTACCTCGAGATGATTGGGATGCTACAGCT 998  
QY 841 TGTGATGTTCTTAACGCTTAAGTCATGTCAGAGTAGAGACAGCCATCAAGTCAAGTAA 900  
Db 999 CGTAGATGTTCTTAACGCTTAAGTCATGTCAGAGTAGAGACAGCCATCAAGTCAAGTAA 1058

QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGCACTGGGATCACTATGAA 960  
Db 1059 CGTTTCATTTGTTACCAACCAAGCTGTGGAAACCAATGCTAGCGGGTCTGCTGATAA 1118  
QY 961 ACCGATCAACAGATATGCTCGCTTGAGATAGCGGTCTACGATTCACATGCTACTAT 1020  
Db 1119 ACCAATCAACAGATATGAGCTGCGCATAGAAATAGGTTATGATTCACATGCAACTAT 1178  
QY 1021 CGTTCACAGTACGCTTAACTACCTCCGAGCAAAATACCTGTTTAACTGGGATAG 1080  
Db 1179 GGTATATATATATCTTATTAATATCCGGAACAAAGCACACTGCTCACTACCTGGATAG 1238  
QY 1081 GGGCTACTTGGCAACATACCTTGCATATACGCCCTGCAACATGCGGAATGCCAGC 1140  
Db 1239 AGCATATTTCCGTTAATATATTAATCTGCAATATGCTTCAACAAATGCAAAATGTCAGC 1298  
QY 1141 CGTGAACCTGTAACAAGATGAGGACTCAACCGGCGCAAGCTTCTTAATGGTCTAGG 1200  
Db 1299 CGTTGAGACTTGTGAATAAGTGCGCTAGATAGAGCTAAACCTTCTTAATGGTCTTGG 1358  
QY 1201 AATCGACTACCCAGATATTAATCAATGCAATTCAGTAAACAAACGGAATCAGA 1260  
Db 1359 TATGAGATATCAAGCATGATATATGCAAAAGCCATTTCAAGTATATACAAAGATCTAA 1418  
QY 1261 CAAAAATATGAGAGCAAGTATGAAAAAGATGCTGCTGCTTACGCTGCTTGGCAATG 1320  
Db 1419 TAAACAAATACGAGCAAGTATGAAAAAATGCTGCTGCTTATGCTTGGCAAAATG 1478  
QY 1321 TGGAACCTACTATTAACCAATGATATTCATTAAGTGTCTTATGATGAGGACGTGAAA 1380  
Db 1479 TGGAACCTACTATTAACCAATGATATTCATTAAGTGTCTTATGATGAGGACGTGAAA 1538  
QY 1381 AGAGTTCTCTAATGTCGGAATCTGCGCATGATGAGAAACGACACCTATATGATAGCGA 1440  
Db 1539 AGAGTTCTCTAATGTCGGAATCTGCGCATGATGAGAAACGACACCTATATGATAGCGA 1598  
QY 1441 CATGATGAAAAACGCTTGTGACTATATGAACTGGAAGAAATGCTTATGCTTGGCTCCC 1500  
Db 1599 CATGATGAAAAACGCTTGTGACTATATGAACTGGAAGAAATGCTTATGCTTGGCTCCC 1658  
QY 1501 TCAGGCTGCTAAGAACAGAACCTTAATCTATACAGACAGAGAAATGAAACCAATCAA 1560  
Db 1659 TCAGGCTGCTAAGAACAGAACCTTAATCTATACAGATGAGAGAAATGAAACCAATCAA 1718  
QY 1561 GACCTGCTAATTTGTAGACCTGATGAACATTTGCTGCGCTATACGCGTAAATTTCAAT 1620  
Db 1719 GAACACTGCTATATGATAGCTCCAGATGAATGTTGTTGCTTATATCTGTAATGTTCTAT 1778  
QY 1621 GCGTGTATGACAGGCTATTTCTAACCGCTGACACACACTTGTAGGCAATGGCCTTACGCT 1680  
Db 1779 GCGTGTATGACAGGCTATTTCTAACCGCTGATTAACCTGTTGAGAGATGGTTTCTAGT 1838  
QY 1681 CGCTGCCAAATTTTACCGGCTATGATGAGACCTACTGCTGTAAGAGCAATCCAGAGA 1740  
Db 1839 TGCAGCTTAAAGTTTATGCTCTATGATGAGACCTACTGCTGTAAGAGCAATCCAGAGA 1898  
QY 1741 TTGCAATATACCAAGGCGCTCTACAGAAATGAGAAATTTCTATTTAAATGCTGCTCG 1800  
Db 1899 TTGCAATATACCAAGGCGCTCTACAGAAATGAGAAATTTCTATTTAAATGCTGCTCG 1958  
QY 1801 TTCTAGCGTGAAGTACCTGCTCCACAACAACCCCATCAACCTGAAGTTCAAGCTCATC 1860  
Db 1959 TTCTAGCGTGAAGTACCTGCTCCACAACAACCCCATCAACCTGAAGTTCAAGCTCATC 2018  
QY 1861 ATCAGATAGTCACTTACAGCTTACTACTCAACCAAGTCAAGCAATATATATAGTACGAC 1920  
Db 2019 ATCAGATAGTCACTTACAGCTTACTACTCAACCAAGTCAAGCAATATATATAGTACGAC 2078  
QY 1921 TACCAATCTTAACCAATATATAGCAACCAATCAACCAACCTGATCAACAAATTCAGAA 1980  
Db 2079 TACCAATCTTAACCAATATATAGCAACCAATCAACCAACCTGATCAACAAATTCAGAA 2138  
QY 1981 TCTTCAACCAAGCACAACCA 1999



|||||  
Db 2139 TCCTCAGCAGCACAACCA 2157  
RESULT 14  
AF139886 2160 bp DNA linear BCT 26-AUG-1999  
LOCUS Streptococcus pneumoniae strain FO-342 penicillin-binding protein  
1a (pbpla) gene, complete cds.  
ACCESSION AF139886  
VERSION AF139886.1 GI:5410462  
KEYWORDS Streptococcus pneumoniae.  
SOURCE Streptococcus pneumoniae.  
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
Streptococcus.  
REFERENCE 1 (bases 1 to 2160)  
AUTHORS Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.  
TITLE Serotype 14 variants of the Spanish penicillin-resistant serotype  
9v clone of Streptococcus pneumoniae arose by large recombinational  
replacements of the cpsA-pbpla region  
JOURNAL Microbiology 145 (Pt 8), 2023-2031 (1999)  
MEDLINE 99392464  
PUBMED 10463168  
REFERENCE 2 (bases 1 to 2160)  
AUTHORS Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.  
TITLE Direct Submission  
JOURNAL Submitted (31-MAR-1999) The Wellcome Trust Centre for the  
Epidemiology of Infectious Disease, Department of Zoology,  
University of Oxford, South Parks Road, Oxford, Oxfordshire OX1  
3PS, UK  
FEATURES  
source Location/Qualifiers  
1. 2160  
/organism="Streptococcus pneumoniae"  
/strain="FO-342"  
/serotype="14"  
/db\_xref="taxon:1313"  
/note="Polish variant of the major penicillin-resistant  
Spanish serotype 9v clone"  
1. 2160  
/gene="pbpla"  
1. 2160  
/gene="pbpla"  
/function="peptidoglycan transglycosylase-transpeptidase"  
/codon\_start=1  
/transl\_table=11  
/product="penicillin-binding protein 1a"  
/protein\_id="AAD43069.1"  
/db\_xref="GI:5410463"  
/translation="MNKPTILRLKYLISFLSLVIAIVLGGVFEYVSKAPSLSE  
SKIVATSSKTYDNKNQILADLSRRRYNAANDPMLVKAIVSIEDHRRFDHIGID  
SIRLIGATRLNLSNLSGGSTLTQOILKLFSTSDQTSRAQDAMLAIOLEOK  
ATKOEILYIVINKYMSNGYGMQIAONYGKIDNNLSLQALALAGMPAPNDYD  
YSHPEAQRDRNLVISEMKNQYISAEQYEAIVPPIIDGQSLKASANYAYMNDYL  
KEVINQVQETGYNLTGMDVYTVNVDQADKHLMDIYNSQOYVSPDDDOVASTAV  
DVSNGKVIADLGARHOASNGTMAQVETNRDMSAKPIIDVAPALIECVYSTATV  
MVDIPYVPCSTSPYVWMDRAYRGNITLQALDQSRVTAIVETLNKGLDRKFFLN  
GLGIDPEMHTANALISNTESNKOYGASSSKMAAIVAFANGGYIYPMYIHKVFS  
DGSKRKESNVGTRAKETJAYMTMDMKTVLTGTGRAYLPMLEQAKTGSTNTDE  
EVENHIKNTGYVADPEMFVYTRKYSMAVMYGSRLPIVGDGLVAAKYRSMY  
LSEGSNPEDMNIPEGLYRNGEFVFKNGARSTWSPAPQOPSTESSSSSDSTSS  
STPSTNSTTNNNNQOSNTPPDOONQNPQAP"  
BASE COUNT 693 a 494 c 407 g 566 t  
ORIGIN  
Query Match 85.3%; Score 1704.6; DB 1; Length 2160;  
Best Local Similarity 90.8%; Pred. No. 0;  
Matches 1815; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

QY 61 TCCCCAGCTAATGATATTTCCACAGATTTGGTTAAGCAATGTTTATCGAAGACCA 120  
Db 219 TCCCCAGCTAATGATATTTCCACAGATTTGGTTAAGCAATGTTTATCGAAGACCA 278  
QY 121 TCGCTTCTTCGACACACAGGGGGATGATACATCCGTATCTGGAGCTTCTGGCAA 180  
Db 279 TCGCTTCTTCGACACACAGGGGGATGATATTCATCCGTATCTGGAGCTTCTGGCAA 338  
QY 181 TCTGCAAGCAATTTCCCTCCAGGTGGATGATCACTCCACCCAAAGTTGATTAAGTGC 240  
Db 339 TCTGCAAGCAATTTCCCTCCAGGTGGATGATCACTCCACCCAAAGTTGATTAAGTGC 398  
QY 241 TTAATTTTCAATTCGACTTCCGACAGACTTTTCTCTAAGGCTCAGGAAGCTTGGTT 300  
Db 399 TTAATTTTCAATTCGACTTCCGACAGACTTTTCTCTAAGGCTCAGGAAGCTTGGTT 458  
QY 301 AGCGATTCTAGTTAGAACAAACCAACCAAGCAAGAAATCTTGACCTACTATATAATTA 360  
Db 459 AGCGATTCTAGTTAGAACAAACCAACCAAGCAAGAAATCTTGACCTACTATATAATTA 518  
QY 361 GGTCTACATGCTATATGGAACTATGGAATGCAGACAGAGCTCAAACTACTATGTAA 420  
Db 519 GGTCTACATGCTATATGGAACTATGGAATGCAGACAGAGCTCAAACTACTATGTAA 578  
QY 421 AGACCCGATTAATTAATTAAGTTTACCTCAGTTAGCTTGGTGGGAATGCTCAGGACAC 480  
Db 579 AGACCCGATTAATTAATTAAGTTTACCTCAGTTAGCTTGGTGGGAATGCTCAGGACAC 638  
QY 481 AAACCAATATGACCCCTATTTCAATCCAGACAGCCCAAGACCCGCAAACTTGGTCTT 540  
Db 639 AAACCAATATGACCCCTATTTCAATCCAGACAGCCCAAGACCCGCAAACTTGGTCTT 698  
QY 541 ATCTGAATGAAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCACTCAATAC 600  
Db 699 ATCTGAATGAAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCACTCAATAC 758  
QY 601 ACCAATTAATGATGAGCTCAAGAGTCTCAAAAGTCAAGCAAGTAATTAATTAATTAAT 660  
Db 759 ACCAATTAATGATGAGCTCAAGAGTCTCAAAAGTCAAGCAAGTAATTAATTAATTAAT 818  
QY 661 TAATTAACCTCAAGAGAGTCAATCAATCAAGTGAAGAAAGCAAGGCTTAATCACTAC 720  
Db 819 TAATTAACCTCAAGAGAGTCAATCAATCAAGTGAAGAAAGCAAGGCTTAATCACTAC 878  
QY 721 AACTGGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780  
Db 879 TACTGGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 938  
QY 781 TTACAATACAGAGAGATAGCTTGGCTATCCACAGCATGAAATGCAAGTGGCTTCAACAT 840  
Db 939 TTACAATACAGAGATAGCTTGGCTATCCCTACAGTGAATGTAATGCAATCTACGGT 998  
QY 841 TGGTGTGCTTTTCAACGTAAGTCAATTCACAGTACAGCAGCCATCACTCAAGTAA 900  
Db 999 GCTAGATGTTTCAAAATGTAAGTCAATTCACAGTACAGCAGCCATCACTCAAGTAA 1058  
QY 901 TGTTCCTTGGAAATTAACCAAGAGTAGAACAACCGGCACTGGGGATCAACTATGAA 960  
Db 1059 CTTTCATTTGGTACCAACCAAGAGTAGAACAACCGGCACTGGGGATCAACTATGAA 1118  
QY 961 ACCGATACAGAGCTATGCTCGCTTGGAGAGAGGTGCTCGATGCAAGTCACTACTAT 1020  
Db 1119 ACCGATACAGAGCTATGCTCGCTTGGAGAGAGGTGCTCGATGCAAGTCACTACTAT 1178  
QY 1021 CTTTCAGATGAGCCCTATTAATCACTTGGAGCAAAATCACTCTGTTTATTAATGGGATAG 1080  
Db 1179 GGTTAATGATATTCCTTAATCACTTGGAGCAAAATCACTCTGTTTATTAATGGGATAG 1238  
QY 1081 GGGCTACTTGGCAATCACTTGGAGCAAAATCACTCTGTTTATTAATGGGATAG 1140  
Db 1239 ACCAATATTCGCTAATTAATCACTTGGAGCAAAATCACTCTGTTTATTAATGGGATAG 1298

QY	1141	CGGGAACCTCTAAACCAAGGTGGACTACACCGCGCCAAGACTTTCCTTAATNGTCTAGG	1200
Db	1299	CGTTGAGACTTTGAATTAAGTTCGGTCTAGATAGACTTAAACCTTCCTTAAGGTCCTGG	1358
OY	1201	AATGACATACCCAACTATTCTACTACTCAAAATGCCATTTCACGTAAACACACCAAGTAACGA	1260
Db	1359	TATGCACATATCCACATGATCATTTATGCAACAGCCATTTCGAATTAATACACAGAAATCTTA	1418
OY	1261	CAAAAAATATGAGCAAGTAGTGTGAAAGATGGCTGCTGCTTACGGTGCCTTTGCAAAATGG	1320
Db	1419	TAAACAATATCGGAGCAAGTAGTAAGTGAATAATGGCTGCTGCTGTTATGCGCTTTGCAAAATGG	1478
OY	1321	TGGAACTACTATTAACCAATGTATATCATTAAGTCTGCTTATGATGGAGTGAANA	1380
Db	1479	TGGCACTACTATTAACCAATGTATATCATTAAGTCTGCTTACAGTGAATGGAATGAANA	1538
OY	1361	AGAGTTCCTAATGTCGGAACCTGCGCATGAGGAANACGACGCTATATGATACCGA	1440
Db	1539	AGAGTTCCTAATGTCGGAACCTGCGCATGAGGAANACGACGCTATATGATACCGA	1598
OY	1441	CATGATGAAAACAGTCTTGACTTATGAAACTGAGAGAAATGGCTTATCTGCTGGCTCC	1500
Db	1599	CATGATGAAAACAGTCTTGACTTATGAAACTGGGCGTGAAGGCTATCTTCTTGCGCTCC	1658
OY	1501	TCAGGCTGGTAAACAGGAACTCTACTATACAGACGAGAAATTTGAAAACACATCA	1560
Db	1659	TCAACTCGTAAACAGGAACTCTACTATACAGATGAGGAATTTGAAAACACATCA	1718
OY	1561	GACCTTCGAATTTGTAGCACTGATGAACCTATTTCGTGGCTTATACGCGTAATATTCAAT	1620
Db	1719	GAACTCTGGCTATGTAGCTCCAGATGAATGTTTGTGTACTCTGAATATTCTAT	1778
OY	1621	GGCCTATGAGCAGGCTATTCTTAACCGCTGACACACTTGTAGGCAATGGCCTTACGCT	1680
Db	1779	GGCTGTATGAGCAGGCTATTGTGAATCGTTTAACTCCTATCGTTGGAGATGCTTCCCTAGT	1838
OY	1681	CGCTGCCAAAGTTTACCGCTCTATGATGACTTACTGCTGTGAAGAGCAATCCAGAGA	1740
Db	1839	TGCACCTTAAAGTTTATTCCTCTATGATGACCTACCTGCTGTGAAGAGCAATCCAGAGA	1898
OY	1741	TTTGAATATACAGAGGGGCTCTACAGAAATGGCAATTCGATTTTAAAAATGGTGCTCG	1800
Db	1899	TTTGAATATATCCAGAGGGGCTCTACAGAAATGGCAATTCGATTTTAAAAATGGTGCTCG	1958
OY	1801	TTTCTACGTGGAACCTACCTGCTCCACAACACCCCATCACTGTAAGTTCACAGCTCATC	1860
Db	1959	TTTCTACGTGGAAGCTACCTGCTCCACAACACCCCATCACTGTAAGTTCACAGCTCATC	2018
OY	1861	ATCAGATAGTTCAACTTACAGACTTAGCTCAACCACTCCAAAGCACAATAATATGATGAC	1920
Db	2019	ATCAGATAGTTCAACTTACAGACTTAGCTCAACCACTCCAAAGCACAATAATATGATGAC	2078
OY	1921	TACCAATCCTTAACATATATAGCGAACATCAAAATCAACCCCTGATCAACAATAATCAGAA	1980
Db	2079	TACCAATCCTTAACATATATAGCGAACATCAAAATCAACCCCTGATCAACAATAATCAGAA	2138
OY	1981	TCCTCAACGACACACAACA 1999	
Db	2139	TCCTCAACGACACACAACA 2157	
RESULT 15			
LOCUS	AF139887	2160 bp	DNA linear BCT 26-AUG-1999
DEFINITION	Streptococcus pneumoniae strain M134 penicillin-binding protein 1a (pbpl1) gene, complete cds.		
ACCESSION	AF139887		
VERSION	AF139887.1	GI:5410464	
KEYWORDS	Streptococcus pneumoniae. Streptococcus pneumoniae Bacteria; Filimicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.		
ORGANISM			

REFERENCE	1 (bases 1 to 2160)
AUTHORS	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.C.
TITLE	Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of Streptococcus pneumoniae arose by large recombinational replacements of the cpsA- <i>pbpA</i> region
JOURNAL	Microbiology 145 (Pt 8), 2023-2031 (1999)
MEDLINE	99392464
PUBMED	10463168
REFERENCE	2 (bases 1 to 2160)
AUTHORS	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.C.
TITLE	Direct Submission
JOURNAL	Submitted (31-MAR-1999) The Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford, Oxfordshire OX1 3PS, UK
FEATURES	Location/Qualifiers
source	1..2160
	/organism="Streptococcus pneumoniae"
	/strain="M134"
	/serotype="14"
	/db_xref="taxon:1313"
	/note="Danish variant of the major penicillin-resistant Spanish serotype 9V clone"
gene	1..2160
	/gene="pbpA"
CDS	1..2160
	/gene="pbpA"
	/function="penicidoglycan transglycosylase-transpeptidase"
	/codon_start=1
	/product="penicillin-binding protein 1a"
	/protein_id="A043070.1"
	/db_xref="GI:5410465"
	/translation="MNKPTLRILIKYISFSLVIAIVLGGVEFFEYVSKAPLSE SKVATFSSIKONKKNLADIAGSERVNAQNDIPDAIKVAISIEDHFPDRGID SIRLQAFKRLNLSNLSGGSTLQOLIKITTSSTSDTIRKQNALAQLOK ATQKELLITVINKYMSNGVGMQTAQNTYGDNLNLSLPQALAGMOPANDYP VSHENAOADRNLVLSEMKNGQYISAQYKAVANTPTDLSVKSASNPAYMDVL KEVSNQVEQETGMLTTGMDVYTNVDOEKHLMIDYNSDQVSPDDLPAASTAV DVSNGKVIADGARHOSANVFCNGMVAETNRPMGSMKP1ITDAPAIEGVYDSTAT MVNDIPVNYGTSPTVNMNPRAVFGNTLTOLALDQSNNTAVETLKKVGLDRKFTPLN GLGIDIPSHIYANNAISSNTTESKQYASSEKMAAATAAFRANGTYKKPIYHKVFS DGSKKFSNVGTRAMKETTYAVMTDMKTVLTITGCGATLPMWPAQKIGISNTDE EVELHKIKNTGYVAPDEMFEVGYTRKYSAAVWTGYSNRLPTLVGGDFLVAKYVSMITTY LSEDTPEHDMTMDGLFRNCEFEVKAFKARNTPMSAPQPPSTRESSSSSSDSSTOSS STPESTNNSSTTTNNNTTQOSNTTPDOONQNPQAP"
BASE COUNT	696 a 495 c 402 g 567 t
ORIGIN	
Query Match	83.9% Score 1677.4; DB 1; Length 2160;
Best Local Similarity	89.9% Pred. No. 0;
Matches 1798; Conservative	0; Mismatches 201; Indels 0; Gaps 0;
1	TAATACTAGACATAAAAAATCAACTCATCTATTGCTGACTTGGGTTCTGAGCCGCGTCAA 60
159	TAAATCTAGACACATAAATAATCAACTCATCTATTGCTGACTTGGGTTCTGAGCCGCGTCAA 218
61	TGCCCAAGCTATGATATATTCCTCCACAGATTTGGTTAAAGGCAATCGTTTCTATCGAAGACA 120
219	TGCCCAAGCTATGATATATTCCTCCACAGATTTGGTTAAAGGCAATCGTTTCTATCGAAGACA 278
121	TGCGTCTTTCGACACAGGGGGATTTGATACATCCGATATCCCTGGGACCTTTCGAGCAA 180
279	TGCGTCTTTCGACACAGGGGGATTTGATACATCCGATATCCCTGGGACCTTTCGAGCAA 338
181	TCTGCAAGCAATTCCTCCACAGGTGATCAACTCTACACCAAGTTGATTAAATTGAC 240
339	TCTGCAAGCAATTCCTCCACAGGTGATCAACTCTACACCAAGTTGATTAAATTGAC 398
241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTGTT 300
399	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTGTT 458

QY 301 AGCGATTCAGTTAGAACAAAAGCAACGAAGAAATCTTGACCTACTATATAATAA 360  
 DB 459 AGCGATTCAGTTAGAACAAAAGCAACGAAGAAATCTTGACCTACTATATAATAA 518  
 QY 361 GGTCTACATGCTTAATGGGAACTATGGAATGAGACAGAGCTCAAACTACTATGTA 420  
 DB 519 GGTCTACATGCTTAATGGGAACTATGGAATGAGACAGAGCTCAAACTACTATGTA 578  
 QY 421 AGACCTCAATTAATTAATGTTACCTCAGTTAGGCTTGCTGGCTGGAATGCTCAGCACC 480  
 DB 579 AGACCTCAATTAATTAATGTTACCTCAGTTAGGCTTGCTGGCTGGAATGCTCAGCACC 638  
 QY 481 AAACCAATATGACCCCTATTCATCCAGACAGCCCAAGACCCGCCAACTGGTCTT 540  
 DB 639 AAACCAATATGACCCCTATTCATCCAGACAGCCCAAGACCCGCCAACTGGTCTT 698  
 QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGTCGACAGATGAGAAAGCAGTCAATAC 600  
 DB 699 ATCTGAATGAAAAATCAAGGCTACATCTCTGTCGACAGATGAGAAAGCAGTCAATAC 758  
 QY 601 ACCAATTAATGAGCTGACCTCAAAAGTCTCAATCAGCAATTAATACCTGCTTACATGGA 660  
 DB 759 ACCAATTAATGAGCTGACCTCAAAAGTCTCAATCAGCAATTAATACCTGCTTACATGGA 818  
 QY 661 TAATTAACCTCAAGGAAGTCAATCAAGTGAAGAAAGAGGCTATTAACCTACTCAC 720  
 DB 819 TAATTAACCTCAAGGAAGTCAATCAAGTGAAGAAAGAGGCTATTAACCTACTCAC 878  
 QY 721 AACTGGATGATGCTCTACACAATGTAGACAGAGAGCTCAAAAACATCTGTGGATAT 780  
 DB 879 TACTGGATGATGCTCTACACAATGTAGACAGAGAGCTCAAAAACATCTGTGGATAT 938  
 QY 781 TTACAATACAGACGATACGTTGCTCTATCCAGAGATGAATTGCAAGTCTTCTACAT 840  
 DB 939 CTACAACTCCGATCAATACGTTGCTCTATCCAGAGATGAATTGCAAGTCTTCTACAT 998  
 QY 841 TTTTGAATGTTTCAAGGATGATGCTGCTTGGAGTATGCTAGATCAATGCTACTAT 900  
 DB 999 CTTAGATGTTTCAAGGATGATGCTGCTTGGAGTATGCTAGATCAATGCTACTAT 1058  
 QY 901 TTTTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGAGTGGGATCAACTATGAA 960  
 DB 1059 GCTTTCATTTGGTACCAACCAAGCTGTGGAACAACATCTGATGGGGTTCCTATGAA 1118  
 QY 961 ACCGATCAGACTATGCTGCTGCTTGGAGTATGCTAGATCAATGCTACTAT 1020  
 DB 1119 ACCGATCAGACTATGCTGCTGCTTGGAGTATGCTAGATCAATGCTACTAT 1178  
 QY 1021 CTTTCAAGATGAGCCCTATAATACCTGGGACAATACTCCTGTTTATTAACGGGATAG 1080  
 DB 1179 GCTTCAAGATGAGCCCTATAATACCTGGGACAATACTCCTGTTTATTAACGGGATAG 1238  
 QY 1081 GGGCTACTTTGGCAACATCACTTGCATATACGCCCTGCAACAATGCGAAACGCTCCAGC 1140  
 DB 1239 AGCATATTTGGTAAATATATCTGCAATATGCTCTTCAACAATGCGAAATGCTCAGC 1298  
 QY 1141 CGTGAACCTTAACCAAGGCTGAGCTCAACCGCGCAAGACTTTCCTAATGCTTAGG 1200  
 DB 1299 CGTGAACCTTAACCAAGGCTGAGCTCAACCGCGCAAGACTTTCCTAATGCTTAGG 1358  
 QY 1201 AATGCACTACCAAGATATCACTACTCAAAATGCCATTTCAAGTAACACAACGGAATCAGA 1260  
 DB 1359 TATGCACTATTCAGATATCACTACTCAAAATGCCATTTCAAGTAATCAACAGAAATCTAA 1418  
 QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAGATGGTGGCTGCTTACGCTGCTTGGCAATAGG 1320  
 DB 1419 TAAACAATATGAGCAAGTAGTGAAGAGATGGTGGCTGCTTACGCTGCTTGGCAATAGG 1478  
 QY 1321 TGGAACTTAATTAACCAATGATATATCAATAAGTCTGCTTATGATGAGGAGTAAAA 1380  
 DB 1479 TGGCACTTACTATTAACCAATGATATATCAATAAGTCTGCTTATGATGAGGAGTAAAA 1538

QY 1381 AGAGTTCTCTAATGTCGAACTGCTGCCATGAGGAAGAACGACAGCTATATGATGACCGA 1440  
 DB 1539 AGAGTTCTCTAATGTCGAACTGCTGCCATGAGGAAGAACGACAGCTATATGATGACCGA 1598  
 QY 1441 CATGATGAAAAACAGTCTTGAATGGAACCTGACGAAATGCCATCTTGTGGCTGCC 1500  
 DB 1599 CATGATGAAAAACAGTCTTGAATGGAACCTGACGAAATGCCATCTTGTGGCTGCC 1658  
 QY 1501 TCAGGCTGGTAAAAACGAAACCTCTAATATACAGACAGAGAAATGAAAAACCATCA 1560  
 DB 1659 TCAGGCTGGTAAAAACGAAACCTCTAATATACAGATGAGAAAGTGAACCAATCA 1718  
 QY 1561 GACCTCTCAATTTGTAGCACTGATGACATTAATGCTGGCTATGACGCTAATATTCAT 1620  
 DB 1719 GACACTGGCTATGTAGCTCCAGATGAAATGTTTGGTTAGTACTGTAAGTATTCAT 1778  
 QY 1621 GGGCTATGAGACAGGCTATTCATACCGTCTGACACCACTTGAAGCAATGGCTTACGGT 1680  
 DB 1779 GGGCTATGAGACAGGCTATTCATACCGTCTGACACCACTTGAAGCAATGGCTTACGGT 1838  
 QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTTACCTGCTGAAGAGAACATCAAGAA 1740  
 DB 1839 TGCAGCTAAAGTTTATCGCTCAATGATTAACGTATCTATGAGAAAGATACATCAGAA 1898  
 QY 1741 TTGGAATATACAGAGGGGCTTACAGAAATGAGAAATGCTTAAAAATGGTCTCG 1800  
 DB 1899 CTGAGAGATGCCAGACGAGCTTTTCAGAAATGAGAAATGCTTAAAAATGGTCTCG 1958  
 QY 1801 TTCTACGTGGAACCTACCTGCTCCACAGAACCCCATCAACTGAAGTTCAAGCTATC 1860  
 DB 1959 TTCTACGTGGAACCTACCTGCTCCACAGAACCCCATCAACTGAAGTTCAAGCTATC 2018  
 QY 1861 ATCAGATGTTCACTTACAGTCTTACAGTCTTACAGTCTTACAGTCTTACAGTCTTAC 1920  
 DB 2019 ATCAGATGTTCACTTACAGTCTTACAGTCTTACAGTCTTACAGTCTTACAGTCTTAC 2078  
 QY 1921 TACCAATCTTAACCAATATACCAATATCAATTAACCCCTGATCAACAAATCAGAA 1980  
 DB 2079 TACCAATCTTAACCAATATACCAATATCAATTAACCCCTGATCAACAAATCAGAA 2138  
 QY 1981 TCCTCAACGACACAAACA 1999  
 DB 2139 TCCTCAACGACACAAACA 2157

Search completed: August 11, 2002, 03:11:46  
 Job time: 28792 sec

•  
•  
•



XX WPI: 1998-272224/24.  
DR P-PSDB: AAM55063.  
XX  
XX Nucleic acid encoding antigenic peptide(s) from Streptococcus  
F1 pneumoniae - or their epitope-containing fragments, useful in  
PT protective or therapeutic vaccines, and for diagnosis  
XX  
PS Claim 1: Page 48: 118pp: English.

CC The present sequence encodes a protein from *Streptococcus pneumoniae*.  
CC The nucleic acid sequence encoding the *Streptococcus pneumoniae* protein  
CC can be useful in vaccines for inducing protective antibodies against  
CC *Streptococcus pneumoniae*, for treatment or prevention of infection e.g.  
CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid  
CC are used to detect *Streptococcus* infection (by usual hybridisation or  
CC amplification methods), also for isolating *Streptococcus* genes or their  
CC allelic variants. The protein can be used similarly to detect specific  
CC antibodies in standard immunoassays, especially for diagnosing or  
CC monitoring infections. Antibodies which bind the protein are used to  
CC detect corresponding antigens, to purify the protein and for passive  
CC immunisation (optionally coupled to a toxin). Vaccines are administered,  
CC e.g. by injection, orally or through the skin, typically at 0.01-1000  
CC (especially 10-300)  $\mu$ g/ml per dose.

**SQ** Sequence 1999 BP; 656 A; 485 C; 376 G; 482 T; 0 other;

Query Match	100.0%	Score 1999;	DB 19;	Length 1999;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1999; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

OY	1	TAAATCTAGCAGCAATAAAAATCAACTCATTTGCTGACTTGCGTTCTGAACGCCGCGTCAA	60
Db	1	taaatctacgacaaataaaaaatcaactcatctgctacttggtctctgaagccgctcaaa	60
OY	61	TGCCCAACCTAATATATATATCCACAGATTTGGTTAAAGCAATGTTTCTATGGAAGGCA	120
Db	61	tgcccaagctaaatgatatatcccaagatcttggttaaggcaatcgcttctatcgaagacca	120
OY	121	TGCGTTTTCGACACACAGGGGGATTTGATACCATCCGTAATCTCGGGAGCTTTCTGGGCAA	180
Db	121	tcgctctctgcacaaaggggatctggtatacaatccgtaatcttcggagcttctctgcgaa	180
OY	181	TCTGCAAGCAATTTCCCTCCAGGTGATCACTCTCACCAACAGTTGATTAAGTTGAC	240
Db	181	tctgcaagcaaatctccctccaaagtgatcaactctcaaccaaaagtttgatlaagttgac	240
OY	241	TTACTTTTCACTTCGATTCGATCCGACGAGCATTTTCTGTAAGGTCGAGAACCTTGGT	300
Db	241	ttaactttcaacttcgacttcgcacagactatcttcgcgaagctcagaagacttggt	300
OY	301	AGCGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATTAATAA	360
Db	301	agcgattcagttagaacaanaagcaacgaagaactcttgactactataataaa	360
OY	361	GGTCTACATGCTCTAATATGGAACACTTGAATATGCACAGCAGCTCAAAACTACTATGCTAA	420
Db	361	ggtctaacctgctctaattgggaacttggaatctgcagacagcagctcaaaactactatgtaa	420
OY	421	AGACCTCAATATTTAAGTTTACCTCAGTTAGCCTTCTGGCTGGAATGCTCAGACACC	480
Db	421	agacctcaataatlaagtttactctcagttagcctctgcgtcggaaatgctctcgaagcac	480
OY	481	AAACCAATATGACCCCTAATTCATCTCAGAAAGCAGGCCCAAGCCGCGAATCTTGCTTT	540
Db	481	aaaccaatatgacccctcatctacatccagaagcagcccaagaccgcgcgaacttgctt	540
OY	541	ATTCGAATATGAAAAATGAAGGCTCATCTCTGCGAACAATATGAGAAAACATCAATAAC	600
Db	541	atctggaatgaaaaatcaaggctcatcatctctgcgaacagtgatgagaagcagtcataac	600
OY	601	ACCATTTACTGATGAGCTACCAAAAGTCTCAATCAGCAAGTAATATACCCTGTTTACATGA	660

Db	601	accacattactgtaggactacaagctcccaaa tcaagaaagtaattaccgcgttacaaggaa	660
Oy	661	TAATATACCTCAAGAGAACTCATCATCAATCAAGTTGAAGAAGAAACAGGCTATATACCTACTAC	720
Db	661	taattactactcaaggaagtctcatcaaaagcttgtagaagaagaacaggtctataactactcaac	720
Oy	721	AACGTGGATGGATGTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
Db	721	aactggagtgtagtgcttacaacaaatgtagccaaagaagcttcaaaaacatctg tgggatalat	780
Oy	781	TTACAAATACAGACGAATATACGTTGGCTTATCCAGACGATGAAATTGGCAAGTGGCTTCTACCAT	840
Db	781	ttacaataacagacgaaatacgtctgcctatccaagacgataatgtcaagcttgcccttaccat	840
Oy	841	TGTTGATGTTTCTACAGGTAAAGCATATGGCCAGCTAGAGACACGGCATCATCTCAAGTAA	900
Db	841	tgttgatgttctctaaagtgtaagtcatctgcccagctaggagacgcatcagtcaagtaaa	900
Oy	901	TGTTTCTCTGGGAATTAACCAAGCAGTAGAAMAACAACCGCATGGGGATCAACTATGAA	960
Db	901	tgtttccctctggaatttaaccaagagtagaanaacaacgcgactgggag tcaactatgaa	960
Oy	961	ACCGATCACAGACTATGCTCTGCTTGGAAGTACGGTGTCTACGATTCACACTGCTACTAT	1020
Db	961	accgatacacgactatgctctcgtcgttgagtagtcggttctacgaltcaactcgtccatct	1020
Oy	1021	CGTTACAGATGATGAGCCCTATATACCTACCCGTGGAGCAAAATACCTCCGTTTATATACGTGGATAG	1080
Db	1021	cgttacaagatagagccctataactataccctcggagacaatactcctglttataactggtgaatag	1080
Oy	1081	GGGCTACTTTGGGCAACATCACTTCGTAATACGGCCCTGCAACAATGCGAAGCTGCCAGC	1140
Db	1081	gggctactcttgtagcaacatactcaacttgtaataagccctgcaacaactcggaaacgtctccagc	1140
Oy	1141	CGTGGAACTCTAAACAAGGTGCGACTCAACCGCCCAAGACTTTTCTTAATGCTCTAGG	1200
Db	1141	cgttgaaactctaaacaaggctcgactccaacgcgcaagacttctctaaatggtcttagg	1200
Oy	1201	AATGACATACCCAAAGTATTCACATACACAAATGCGCATTTCCAAAGTAAACAACAACCGATTCAGA	1260
Db	1201	aatcgataccaacgaagtatctcaactatcaaatgcatcttaagvlaaacaacacgaaatcaga	1260
Oy	1261	CAAAAAATATGAGACAGATAGTAAAGAAAGAGCGCTGCTTACGCGTCTTTGGCAAAATCG	1320
Db	1261	caaaaaatatgtagcagaagttagaaaaagatgagctgctgttaagctgcctcttgcaaatagg	1320
Oy	1321	TGGAACCTTACTATTAACCAATGTATATCCATTAAGTCGCTTTAGTGATGGAGTGAATAA	1380
Db	1321	tggaacttactataaacaatgtaatacctaaagctaaagctctttagtgaatgtaggagtagaaaa	1380
Oy	1381	AGACTTCTCAATGTCGAGCATCGTGCACATGMAAGMAAGACAGACGCTATATGATGACGGA	1440
Db	1381	agacttctctaaatgctggaactcgtgcacatggaaggaacaaagcctatataatgtagccga	1440
Oy	1441	CATGATGAAAAACATCTTGGACTTATGGAAGCTGACCAAAATGCCATATCTTGGCTTGGCTCC	1500
Db	1441	catgatagaaaaacgctcttgactataatggaacttgagcgaataatgcttatcttgcttgctccc	1500
Oy	1501	TCAGGCTGGTAAACAGGAACCTCTTAACATACAGACGAGAAATTTGAAACACATCAAA	1560
Db	1501	tcaagctcgttaaaaacaggaacctctaactatacagaacgagaaatltgaaaacacatcaaa	1560
Oy	1561	GACCTCTCAATTTGTAGCACCTGATGATCACTATTTGCTGGCTTACGCGTAAATATTCAAT	1620
Db	1561	gacctctccaatttgtagcacctgtagaactatttgctgtctatacgcgtaaatattcaat	1620
Oy	1621	GGCTGTATGACAGGCTATCTTAACCGTGTGACACACACTTGTAGCAATGCGCTTACCGT	1680
Db	1621	ggctgtatgacagagctatctctaaacgctctgaaacacactttagagcaatgagccttaaggtc	1680
Oy	1681	CGGTGCAAGATTATACGCTCTATGATGACCTACCTGTCTGAAGGAACCAATCCAGAGA	1740

Db	1681	cgctgccaaagtttacgcgtctatgatgacctactcgtctgtgaaggaagcaatccagaaga	1740
Oy	1741	TTGGATATATACCAAGAGGGGCTCTACAGAAATGGAGAAATTCGTATTTTAAATAGTGCTCG	1800
Db	1741	ttggatatataccagaagggtctctacagaatggagatctgtatltttaaataigtgctcg	1800
Oy	1801	TTCTACAGTGGAACTACCTGCTCCACACACACACCCCATCATCACTGAAGAATTCAAGCTCATC	1860
Db	1801	ttctacagtggaactcacctcgtctcaacaacaccccatcaactgaaagtccaagctcatc	1860
Oy	1861	ATCAGATAGTTCACACTCTACAGCTGTAGCTGCATCCACTCCAGACACAATATATAGTACAGC	1920
Db	1861	atccagatagtgtaactcaactcagctcgtctcaaccactccaagaacaataatagtaagac	1920
Oy	1921	TACCAATCCTTAACAATATATACGCAACAATCAAAATPACACCCCTGATCAACAAATCAGAA	1980
Db	1921	taccaatcttaacataataatagcgaacaacataacataacacccctgatacaacaataatcagaa	1980
Oy	1981	TCCTCAACCGACGACACACCA 1999	
Db	1981	tcctcaaccagcacacacca 1999	
RESULT 2			
AAV52278/c			
ID	AAV52278 standard; DNA; 10711 BP.		
XX			
AC	AAV52278;		
XX			
DT	23-OCT-1998 (first entry)		
XX			
DE	Streptococcus pneumoniae genome fragment SEQ ID NO:145.		
XX			
KM	Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;		
KW	computer readable medium; vaccine; pharmaceutical composition; ds.		
XX			
OS	Streptococcus pneumoniae.		
XX			
PN	WO9818931-A2.		
XX			
PD	07-MAY-1998.		
XX			
PF	30-OCT-1997; 97WO-US19588.		
XX			
PR	31-OCT-1996; 96US-0029960.		
XX			
PA	(HUMAN) HUMAN GENOME SCT INC.		
XX			
PI	Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;		
PI	Kunsch CA, Rosen CA;		
XX			
DR	WPI: 1998-272225/24.		
XX			
PT	Computer-readable medium with recorded Streptococcus pneumoniae		
PT	polynucleotide sequences - useful in diagnostic kits and assays, and		
PT	pharmaceutical compositions and vaccines for Streptococcus		
PT	pneumoniae		
XX			
PS	Claim 1: Page 972-978; 1409pp; English.		
XX			
CC	The present invention describes a computer readable medium which has		
CC	the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)		
CC	recorded on it, or a representative fragment or a sequence at least 95%		
CC	identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in		
CC	SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from		
CC	Streptococcus pneumoniae. The present invention also describes an		
CC	isolated nucleic acid molecule encoding a homologue of any of the		
CC	fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the		
CC	nucleic acid molecule is produced by a process comprising: (a) screening		
CC	a genomic DNA library using as a probe a target sequence defined by any		
CC	of the sequences in SEQ ID NO:1 to 391, identifying members of the		
CC	library which contain sequences that hybridise to the target sequence and		
CC	isolating the nucleic acid molecules from the members; or (b) isolating		

Query Match	100.0%	Score 1999;	DB 19;	Length 10711;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1999;	Conservative	0;	Mismatches	0; Indels
				Gaps
				0;
CC	mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid			
CC	molecules whose nucleotide sequence is homologous to amplification			
CC	primers derived from the fragment of the <i>S. pneumoniae</i> genome to prime			
CC	the amplification and isolating the amplified sequences. The computer			
CC	readable medium can be used in a computer-based system for identifying			
CC	fragments of the <i>S. pneumoniae</i> genome of commercial importance, or			
CC	expression modulating fragments of the <i>S. pneumoniae</i> genome. Products			
CC	from the present invention can be used in diagnosis kits and assays, and			
CC	pharmaceutical compositions and vaccines for <i>S. pneumoniae</i> .			
XX				
Sequence 10711 BP; 3280 A; 2056 C; 2412 G; 2963 T; 0 other;				
QY	1 TAAATCTAGCAATAAAAATCACTATGCTGACTGTTGGTTCTGAAGCCGCGCTCAA	60		
DB	9767 TAAATCTGCGCAATAAAAATCACTATGCTGACTGTTGGTTCTGAAGCCGCGCTCAA	9708		
QY	61 TCGCCAGGTATGATTTCCACAGATTTGGTTAAGCAATCGTTTATGGAAGACCA	120		
DB	9707 TCGCCAGGTATGATTTCCACAGATTTGGTTAAGCAATCGTTTATGGAAGACCA	9648		
QY	121 TCGCTTCTCGACACACAGGGGGATGATACATCCGTATCCTGGAGCTTTCTTGGCAA	180		
DB	9647 TCGCTTCTCGACACACAGGGGGATGATACATCCGTATCCTGGAGCTTTCTTGGCAA	9588		
QY	181 TCTGCAGCAATTCCTCTCCAAAGTGTGATCAACTCTCACCCCAAGTTATTAAGTTGAC	240		
DB	9587 TCTGCAGCAATTCCTCTCCAAAGTGTGATCAACTCTCACCCCAAGTTATTAAGTTGAC	9528		
QY	241 TTACTTTTAACTGACTTCGACAGCACTATTTCTGTAAGGCTCAGGAAGCTTGAT	300		
DB	9527 TTACTTTTAACTGACTTCGACAGCACTATTTCTGTAAGGCTCAGGAAGCTTGAT	9468		
QY	301 AGCGATTCAGTTAGAAACAAAAGCAACACAGAGAAGAAATCTTGACTACTATATAATTA	360		
DB	9467 AGCGATTCAGTTAGAAACAAAAGCAACACAGAGAAGAAATCTTGACTACTATATAATTA	9408		
QY	361 GGTCTACATGCTTAATGGCACTATGGAATGCAGACAGCAGCTCAAACTACTATGTTAA	420		
DB	9407 GGTCTACATGCTTAATGGCACTATGGAATGCAGACAGCAGCTCAAACTACTATGTTAA	9348		
QY	421 AAGCCGATTAATTAAGTTTAAGTTTAAGTTTAAGTTTAAGTTTAAGTTTAAGTTTAAGTTTA	480		
DB	9347 AAGCCGATTAATTAAGTTTAAGTTTAAGTTTAAGTTTAAGTTTAAGTTTAAGTTTAAGTTTA	9288		
QY	481 AAACCAATATGACCCCTATTCACATCCAAACAGACGCCCAAGACCGCGCAACTTGGCTTT	540		
DB	9287 AAACCAATATGACCCCTATTCACATCCAAACAGACGCCCAAGACCGCGCAACTTGGCTTT	9228		
QY	541 ATCTGAATGAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC	600		
DB	9227 ATCTGAATGAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC	9168		
QY	601 ACCAATTAATGATGACCTCAAGAGCTCAAAATCAGCAAACTAATTAATCCCGCTTACATGGA	660		
DB	9167 ACCAATTAATGATGACCTCAAGAGCTCAAAATCAGCAAACTAATTAATCCCGCTTACATGGA	9108		
QY	661 TAATTAATCAAGGAAGTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA	720		
DB	9107 TAATTAATCAAGGAAGTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA	9048		
QY	721 AACTGGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAAATCTGTGGATAT	780		
DB	9047 AACTGGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAAATCTGTGGATAT	8988		
QY	781 TTACCAATACAGAGCAATGAGTGGCTATCCACAGACGATGAATTCGCAATGCTGCTTACAT	840		
DB	8987 TTACCAATACAGAGCAATGAGTGGCTATCCACAGACGATGAATTCGCAATGCTGCTTACAT	8928		

```
QY 841 TGTGATGTTTCTAAGCGTAAAGTCATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
    |||
Db 8927 TGTGATGTTTCTAAGCGTAAAGTCATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 8868
QY 901 TGTTCCTTGGAAATTAACCAAGCAGTGAAGAACCAACCGGAGCTGGGGATCAACATGTAA 960
    |||
Db 8867 TGTTCCTTGGAAATTAACCAAGCAGTGAAGAACCAACCGGAGCTGGGGATCAACATGTAA 8808
QY 961 ACCGATCAGACAGTATGCTCCTGCTGGAGTAGGCTGTACGATTCACATTCGCTACTAT 1020
    |||
Db 8807 ACCGATCAGACAGTATGCTCCTGCTGGAGTAGGCTGTACGATTCACATTCGCTACTAT 8748
QY 1021 CGTTCAGATAGAGCCCTTATCTACCTACCTGGAGCAATATCTCTTTATTAACCTGGAGTAG 1080
    |||
Db 8747 CGTTCAGATAGAGCCCTTATCTACCTACCTGGAGCAATATCTCTTTATTAACCTGGAGTAG 8688
QY 1081 GGGGTACTTTGGCAACATCACTTGGCAATAGCCCTGGCAACATCGGCAACGTCGCCAGC 1140
    |||
Db 8687 GGGGTACTTTGGCAACATCACTTGGCAATAGCCCTGGCAACATCGGCAACGTCGCCAGC 8628
QY 1141 CGTGAACACTCTAAACAAGTCGAGCTCAACCGGCGCAAGACTTTCCTAAATGGCTCTAGG 1200
    |||
Db 8627 CGTGAACACTCTAAACAAGTCGAGCTCAACCGGCGCAAGACTTTCCTAAATGGCTCTAGG 8568
QY 1201 AATGCACTACCCAGATATCTACTACTCAATGCCATTTCAAGTAACACACCGAATCAGA 1260
    |||
Db 8567 AATGCACTACCCAGATATCTACTACTCAATGCCATTTCAAGTAACACACCGAATCAGA 8508
QY 1261 CAAATAATATGAGACAGTACTGAAAAAGATGGCTGCTCTTACCGCTTTCGCAAAATGG 1320
    |||
Db 8507 CAAATAATATGAGACAGTACTGAAAAAGATGGCTGCTCTTACCGCTTTCGCAAAATGG 8448
QY 1321 TGAAGTACTATATAACCAATATATCCATTAAGTCTTTAGTGAATGGAGTGAATAA 1380
    |||
Db 8447 TGAAGTACTATATAACCAATATATCCATTAAGTCTTTAGTGAATGGAGTGAATAA 8388
QY 1381 AGAGTTCTATATGTCGAACTGCTGCCATGAGAGAAACGACAGCTATATGATGACCGA 1440
    |||
Db 8387 AGAGTTCTATATGTCGAACTGCTGCCATGAGAGAAACGACAGCTATATGATGACCGA 8328
QY 1441 CATGATGAAACAGTCTGATTAATGAAAGTGAACGAAATCCATCTGTTGGCTGCC 1500
    |||
Db 8327 CATGATGAAACAGTCTGATTAATGAAAGTGAACGAAATCCATCTGTTGGCTGCC 8268
QY 1501 TCAGGCTGTATAAACAGAGAACTCTAATATACAGAGAGAGAAATTTGAAAAACCAATCAA 1560
    |||
Db 8267 TCAGGCTGTATAAACAGAGAACTCTAATATACAGAGAGAGAAATTTGAAAAACCAATCAA 8208
QY 1561 GACCTCTCAATTTGTAGCACTGATGAACATATTTGGTGGCTATACGGCTAAATATTCAT 1620
    |||
Db 8207 GACCTCTCAATTTGTAGCACTGATGAACATATTTGGTGGCTATACGGCTAAATATTCAT 8148
QY 1621 GGGGTATGAGCAGGCTATGTAACGCTGACACCACTGTAAGCAATGGCCTTAGCGT 1680
    |||
Db 8147 GGGGTATGAGCAGGCTATGTAACGCTGACACCACTGTAAGCAATGGCCTTAGCGT 8088
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTTACCTGCTGAAGAGCAATTCAGAGA 1740
    |||
Db 8087 CGCTGCCAAAGTTTACCGCTCTATGATGACCTTACCTGCTGAAGAGCAATTCAGAGA 8028
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAATTCGTATTTAAAAATGGTCTCG 1800
    |||
Db 8027 TTGGAATATACAGAGGGGCTCTACAGAAATGGAATTCGTATTTAAAAATGGTCTCG 7968
QY 1801 TTCTAGCTGGAACACCTGCTGTCACAACACCCCAATCACTGAAGTTCAACCTCATC 1860
    |||
Db 7967 TTCTAGCTGGAACACCTGCTGTCACAACACCCCAATCACTGAAGTTCAACCTCATC 7908
QY 1861 ATCAGATAGTTCAACTACAGTCTAGCTCAACCACTCAAGCAACAATAATATAGTACAG 1920
    |||
Db 7907 ATCAGATAGTTCAACTACAGTCTAGCTCAACCACTCAAGCAACAATAATATAGTACAG 7848
QY 1921 TACCAATCTTAACAATAATACGCAACAATCAATAATCAACCCCTGATCAACCAAAATCAGAA 1980
    |||
```

```
Db 7847 TACCAATCTTAACAATAATACGCAACAATCAATAATCAACCCCTGATCAACAAATCAGAA 7788
    |||
QY 1981 TCCITCAACGACGACACCA 1999
    |||
Db 7787 TCCITCAACGACGACACCA 7769
    |||

RESULT 3
AAT08027
ID AAT08027 standard; DNA; 2048 BP.
XX
AC AAT08027:
XX
DT 03-DEC-1996 (first entry)
XX
DE S. pneumoniae penicillin binding protein 1A soluble variant DNA.
XX
KW Penicillin binding protein; PBP 1A; bifunctional protein;
KW transglycosylase; transpeptidase; identification; assay; inhibitor;
KW antibiotic resistant; bacteria; soluble variant; protein structure;
KW X-ray crystallography; determination; ds.
XX
OS Streptococcus pneumoniae.
XX
FH Key Location/Qualifiers
FT CDS 1..2046
FT /tag= a
FT
PN GB2290792-A.
PD 10-JAN-1996.
XX
PF 29-JUN-1995: 95GB-0013306.
XX
PR 24-NOV-1994: 94SE-0004072.
PR 01-JUL-1994: 94IN-0000580.
XX
PA (ASTR ) ASTRA AB.
XX
PI Balganesch TS, Town CM.
XX
DR WPI: 1996-042232/05.
DR P-PSDB: AAW04359.
XX
PT Sol. derivs. of bifunctional penicillin binding protein (BPBP) -
PT opt. lack transglycosylase activity, useful to identify and assay
PT for antibodies or cpds. which bind BPBP
XX
PS Claim 11: Pages 65-68; 108pp; English.
XX
CC The present sequence encodes a soluble variant of the S. pneumoniae
CC penicillin binding protein (PBP) 1A, which lacks the 1st 38 amino
CC acid residues of the wild type protein. Wild type PBP is a
CC bifunctional protein, which binds the cell membrane when expressed
CC in a bacterial cell, having transglycosylase and transpeptidase
CC activities. The variant protein (NCIMB 40665) in conjunction with
CC a labelled anti-bifunctional PBP monoclonal antibody, can be used
CC to identify and assay for cpds. which bind bifunctional PBP. Such
CC cpds., as inhibitors of bifunctional PBP have a potential use in
CC therapeutic cpds. which inhibit the growth of antibiotic resistant
CC bacteria. The soluble variant may also be used in X-ray
CC crystallography.
XX
SQ Sequence 2048 BP; 667 A; 498 C; 390 G; 493 T; 0 other;
```

Query Match 99.1%; Score 1981.4; DB 17; Length 2048;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1 TAAATCTAGACAAATTAATATCACTGATTCGACTTGGGTTCTGAGCGCCGCGTCAA 60  
|||||



48 taataactacgacaataaatacaatcatctgctgacttggtgtcttgaaagccgcgctcaa 107  
QY 61 TCCCAAGCTATGATATTCCACAGATTGGTTAAGCAATCGTTTCATGGAAGACCA 120  
Db 108 tgcacaagctaatgaaatctccacagacttggttaaggaatcgtttctcatcgaaagaca 167  
QY 121 TCGCTTCTTCGACCAAGGGGGATTGATACATCCGTATCTGGAGAGCTTTTCGCGAA 180  
Db 168 tgcgtctctcgacccaaggggagatgataccatccgtatccctggagagcttcttcgcgcaa 227  
QY 181 TCTGCAAAAGCAATTCCTTCACAGGTGATCAACTCTACCCAAAGTTGATTAACTTGAC 240  
Db 228 tctgcaaaagcaatctccctcccaagtgatcagctctcactcaacagcttgattaaagtgcac 287  
QY 241 TTACTTTTCAATTTGAGCTTCGACAGACTTTTCTCGTAAGGCTCAGGAAGCTTGGTT 300  
Db 288 ttaacttcaactcgaactccgaccagactatctctgtaagagctcaggaagcttggtt 347  
QY 301 AGCGATTGATTAGAAAACAAACGACAGCAAGAAATCTGACCTACTATATAATTA 360  
Db 348 agcgattcaagttagaaacaaaagcaacaaagaaatcttgactactatataataa 407  
QY 361 GGTTCATGTCATTGGAAGACTATGAAATGACACAGCAGCTCAAAACTACTATGTAA 420  
Db 408 ggtctacatgcttaagtggaaactatggaatgacagacagactcaaaactactatgtaa 467  
QY 421 AGACGTCAATATTAAAGTTTAACTGCTAGTTAGCTTGGTGGTGGAAATCCCTCAGGCAC 480  
Db 468 agacctcaataatlaaagtttaacctagttagcttgctggtgtaatgcttcccaagcac 527  
QY 481 AAACCAATATGACCCCTATTTCATCTCAGAACAGCAGCCCAAGCCCGAACTTGTT 540  
Db 528 aaaccaataatgacccctattccatcccgaaagcagcccaagccgcgcgaacttggtctt 587  
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTAAATAC 600  
Db 588 atctgaaatgaaaaatacaaggtacatactctgtgtaacagtaatgaaagcagtaaac 647  
QY 601 ACCAATTCTGATGACTACAAAGTCTCAATCAGCAAGTATTTCCTGCTTAATGGA 660  
Db 648 accaatctcgtatggcttacaagttccaatccaagaaatlaattccctgcttaactgta 707  
QY 661 TAAATACCTCAAGAAATCATCATCAAGTTGAAGAAAGAGGCTATTAACCTACTCAC 720  
Db 708 taataactcaaggaaggtcatcaatccaagttaagaaagaacaggtataaactctacc 767  
QY 721 AACTGGATGATGTCACAAATGTAGACCAAGAGCTCAAAAAACATCTTGAGGATAT 780  
Db 768 aactggatgtagtgtctacacaatgttagaccagaaggtccaataacatctgtggat 827  
QY 781 TTACAAATCAGAGCAATACGTGCTTATCCAGACGATGAATTGCAAGTGGCTTACCAT 840  
Db 828 ttaacaatacagaatcagttcgtctatccagaatgaaatgcaagttcgtcttcaacat 887  
QY 841 TGTGATGTTTCTTAACGGAATGATGATGTCAGCTAGAGCAGCCCATCATCAAGTAA 900  
Db 888 tgttgaatgttcttaacggtlaaagttatgtccagcttagagagacgcatacgtcaagt 947  
QY 901 TGTTCCTTGGGAATTAACCAAGAGTAGAAACCAACCGCATGGGATCAACTATGAA 960  
Db 948 tgttctcttcggaatlaaaccaagcagtagaaacaacgcgcgtggggtcaactatgaa 1007  
QY 961 ACCGATTCACAGACTATGCTCCTGCTTGAGTAGCGTGTCTAGATTCAACTGCTACTAT 1020  
Db 1008 accgatcaagactatgctccctgcttgagtagagtgltctcagagtaacagtcacact 1067  
QY 1021 CGTTACGATGAGCCCTATTAATCACTACCTGGAGCAATACCTGTTTATTAATGGGAATG 1080  
Db 1068 cgttaacagatgagccctataactaacctcgtggagcaataacccctgtttataacttgga 1127  
QY 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAAATGCGAAGCTCCACAC 1140  
Db 1128 gggctactttggcaacatcaacttgcaataagcctcgcaacaatcgcgaaagctccagc 1187

QY 1141 CGTGAACCTTAAACAAGGTGAGCTCAACCGCGCAAGACTTTCCTAAATGTCATAG 1200  
Db 1188 cgtggaaactccaataaagaatcgagatcaacccgcgcgaacttctcctaataatggtctcg 1247  
QY 1201 AATGCACTACCAAGATATTCACTACTCAAAATGCCATTTTCAATTAACACAAACGAATCAGA 1260  
Db 1248 aatgcatacccaagtatactactactcaaatgcatcttcaagttaacacacacccaatcgaa 1307  
QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAAGTGGCTGCTTACGCTTGCCTTTGCAAAATG 1320  
Db 1308 caaaaataatgagcaagtagtgaagaagtgctgctgttactgcgtcccttgcaaatg 1367  
QY 1321 TGGACTTACTATTAACCAATGATATTCATTAAGAGTCGCTTATGATGATGGAGGAA 1380  
Db 1368 tggaaacttactataaaccgaatgatalccataaagtcgtcctttagtgaagggaatgaa 1427  
QY 1381 AGAGTTCTTAAATGTCGGAAGTCTGCCATGGAAGGAACGACAGCCTATATGATGACGA 1440  
Db 1428 agagttctttaaagtcgagactcgtgcataagaaagaaagacagcctataatgatacgca 1487  
QY 1441 CATGATGAAGAAGTCCTTGAATTTATGAACTGACGAAATGCTTATCTTGGCTGCC 1500  
Db 1488 catgataaacaagctcttgatgtaagaaatgagcaaatgctctatcttgcttgctccc 1547  
QY 1501 TCAAGCTGTAACAGAGACCTCAACATACAGACGAGGAATTTGAAACACATCA 1560  
Db 1548 tcaagcttgtaaaaacaggaactccaactatacagaacgaggaatctgaaacacaatcaa 1607  
QY 1561 GACCTCTCAATTTGAGCACTGATGAATATTTGCTGCTGCTTACGCGTAATATTCAAT 1620  
Db 1608 gacctctcaatttgtagcaacctgagaaactattgctgctataagcgttaaatctcaat 1667  
QY 1621 GGCTGATGAGCAGGCTATTCTAACCGTCTGACACCACTTGTAGGCAATGSCCTTACGGT 1680  
Db 1668 ggcgtatgagcaagctatcttaaccgcttgacacacaacttgaaaggaatgyccttaag 1727  
QY 1681 CGCTGCCAAAGTTTACCGCTCATATGATGACCTGCTGTGAAGGAATTCAGGAAGA 1740  
Db 1728 cgcgtccaaagtttaacgctctatgataagttccactcgtcttgaaaggaatccagaa 1787  
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAATAATGTCGTCG 1800  
Db 1788 ttggataataccagggggtcttaacagaatgagaatctgatttaaaaaatggtgctcg 1847  
QY 1801 TTCTAGTGAAGCTACCTGCTCCACAACACCCCATCACTGAAGTTCAAGCTCATC 1860  
Db 1848 tctcagtgaggtcacctgcctccaacaaccccatcaactgaaagtccaagctcac 1907  
QY 1861 ATCAGATAGTTCAACTTCACACTGTACAGTCAACCACTCCAAACGACAAATTAATGATGAC 1920  
Db 1908 atcagatagttcaactccaagttcagctcaacacactccaagcaaaaataatagtagac 1967  
QY 1921 TTACCAATCCTTAATTAATACGCAACAAATCAATCAACCCCTGATTCACAACAAATCAGAA 1980  
Db 1968 taccaatcttaacataataacgcaacaatccaatactcaacacctgattcaacaaaatccagaa 2027  
QY 1981 TCTCTAACCGACACAACCA 1999  
Db 2028 tctctcaaccagacaacca 2046

RESULT 4  
AAH02055  
ID AAH02055 standard; DNA; 2160 BP.  
XX  
AC AAH02055;  
XX  
XX 24-JUL-2001 (first entry)  
XX  
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2048.  
XX  
XX Species specific; genus specific; family specific; probe; detection;  
KW

KW identification; algal; archaeal; bacterial; fungal; parasitica;  
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;  
 KW translation elongation factor G; RecA recombinase; resistance;  
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;  
 KW vaccine; primer; ds.  
 XX  
 OS Streptococcus pneumoniae.  
 PN WO200123604-A2.  
 PD 05-APR-2001.  
 PF 28-SEP-2000; 2000WO-CA01150.  
 PR 28-SEP-1999; 99CA-2283458.  
 PR 19-MAY-2000; 2000CA-2307010.  
 XX  
 PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.  
 PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;  
 PI Picard FJ, Roy PH;  
 XX  
 DR WPI: 2001-24506/25.  
 XX  
 PT Nucleic acid sequences are used to generate universal probes and  
 PT primers which can be used to identify and detect the presence of algal,  
 PT archaeal, bacterial, fungal and parasitica species in a test sample -  
 PS Disclosure; Page 1463-1464; 1580pp; English.  
 XX  
 CC The present invention describes a method for generating a repository of  
 CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes  
 CC and/or primers are derived. The method comprises amplifying the nucleic  
 CC acids of determined algal, archaeal, bacterial, fungal and parasitica  
 CC species with a combination of defined primer pairs. The method can be  
 CC used for producing probes and/or primers for detecting one or more  
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and  
 CC parasites, for universal detection and for specific and ubiquitous  
 CC detection and identification of an algal, archaeal, bacterial, fungal  
 CC and parasitica species, genus, family and group. A nucleic acid (II)  
 CC obtained using the method of the invention can be used for the universal  
 CC detection of any bacterium, fungus or parasite in a sample and for the  
 CC detection of at least one antimicrobial agent resistance gene or at  
 CC least one toxin gene. hexa nucleic acids are used for the specific and  
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.  
 CC (I) can be used to design a therapeutic agent which is effective against  
 CC microorganisms. Microbial species or genus or family or phylum or group  
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,  
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,  
 CC Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,  
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests  
 CC provides faster results than substrate specificity tests as results can  
 CC be determined in an hour and improved accuracy is also achieved.  
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes  
 CC which are given in the exemplification of the present invention.  
 XX  
 SQ Sequence 2160 BP; 697 A; 522 C; 413 G; 528 T; 0 other;

Db 279 tgccttcgcacacaaagggatgataccatccgtaacctggaggtcttcgcgcaa 338  
 Oy 181 TGTGCAAGCAATTCCTCCCAAGGATGATCACTCCCAAGCAAGTGGATTAACTGAC 240  
 Db 339 tctgcaagaacatcccccacaaagtgatcagctctcaacacagtgatgaagtgac 398  
 Oy 241 TTACTTTTCAACTTCGACTTCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300  
 Db 399 ttactttcaacttcgacttcgcacagacatacttctcgtgaagtgctcgaagcttgct 458  
 Oy 301 AGCGATTGATGATGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360  
 Db 459 agcgattgattgatacacaacaaagcaacaaagcaacaaagcaacaaagcaacaaagca 518  
 Oy 361 GGTCTACATGCTTAATGGAAGCAATGATGATGATGATGATGATGATGATGATGATGAT 420  
 Db 519 ggtctacatgcttaatggaagcaatgatacagagcaagcaagcaagcaagcaagcaagca 578  
 Oy 421 AGACCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480  
 Db 579 agacctcaataatlaagttacactgaagttacactgaagttacactgaagttacactga 638  
 Oy 481 AAACCAATATGACCCCTTTTACATCCAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 540  
 Db 639 aaaccaatataccctctatccacatccagagcaagcccaagcaagcccaagcaagccca 698  
 Oy 541 ATCTGAATGAAATTAATCAAGGCTACATCTGCTGCAACAGATGAGAAAGCAAGTATAC 600  
 Db 699 atctgaaatgaaaatacaagagctacatctcgcggaagatgatagaagaagcaataac 758  
 Oy 601 ACCAATTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 Db 759 accaatctgattggtctcaacaaagctcacaatccagcaagcaagcaagcaagcaagca 818  
 Oy 661 TAATTACTGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 720  
 Db 819 taattactctcaagaaagcaagcaagcaagcaagcaagcaagcaagcaagcaagcaagca 878  
 Oy 721 AACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 Db 879 aactggatgattggtctcaacaaagctcacaatccagcaagcaagcaagcaagcaagca 938  
 Oy 781 TTACAAATACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 840  
 Db 939 ttacaatacagaagcaatcgtctcctacacagacagatgaaagcaagcaagcaagcaag 998  
 Oy 841 TGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
 Db 999 tgtgtgattgtctcctacacagcaagcaagcaagcaagcaagcaagcaagcaagcaag 1058  
 Oy 901 TGTTCCTTCGGAATTTAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 960  
 Db 1059 tgttccctcgtgaattacccaagcaagcaagcaagcaagcaagcaagcaagcaagcaag 1118  
 Oy 961 ACCGATACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1020  
 Db 1119 accgatcaagcaagcaagcaagcaagcaagcaagcaagcaagcaagcaagcaagcaag 1178  
 Oy 1021 CGTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 Db 1179 cglttcagatgattggtctcctacacagcaagcaagcaagcaagcaagcaagcaagca 1238  
 Oy 1081 GGGCTACTTGGCAACATCACTTCGCAATAGCCCTGCAACAAATCGGCAAGCAAGTCCAGC 1140  
 Db 1239 gggctacttggcaacatcacttcgcaatagccctcgcaaacatcggaagcaagcgcacagc 1298  
 Oy 1141 CGTGAATACTCTAAACAAGTCTGAGCTCAACCGGCGCAAGCAAGTCTTAAATGGTCTAGG 1200  
 Db 1299 cgtgaaactcttaacaagcgtgagctcaacccgcaagcaagccttcttaaatgctcgg 1358  
 Oy 1201 AATGCACTACCAAGATTTACTACTACTCAATAGCAATTTCAAGTAAACCAACCAAGTACA 1260  
 Db 1359 aatgcactaccagataltactactcaaatagcaatttcaagtaagtaacacacagcaagca 1418

```
QY 1261 CAAAAATATGAGACAGTAGTGAAGAATGCGTCTGCTTACCGCTGCTTGGCAATGG 1320
Db 1419 caaaaalatalgagcaagtagtgaagaatgagctgcttaccgcttgccttgaatg 1478
QY 1321 TGAACATTACTATTAACCAATGTATATCAATAAAGTGTCTTATGATGAGGATGAAA 1380
Db 1479 tggaaactactataacaatagtatatacaataaagtcgtctttagtgatgagtgaaaa 1538
QY 1381 AGAGTTCTCTAATGTCGAACTGCTGCATGAGGAAAAGCAAGCATATATGATGACGA 1440
Db 1539 agagttctcctaattgtaggaactcgtgcatagaagaacgacgacataatgtagcaga 1598
QY 1441 CATATATAAAACAGCTCTTACTTATGAGACGAGCAAAATCCATATCTGTTGGCTCC 1500
Db 1599 calgaatgaacaacagctcttgatgataagcaatgagaaatgacctatcttgctgctcc 1658
QY 1501 TCAGGCTGTGTAACAGGAACTCTTACTATACAGAGGAAATGAAAAACACATCAA 1560
Db 1659 tcagggctgtaaaaaacagaaacctataactatacagaggaatgtgaaaaacacataa 1718
QY 1561 GACCTTCATTTGTAGACACCTGTATGACTATTTGCTGGCTATACGGCTAAATATCAAT 1620
Db 1719 gacctcccaattgtagcacctgtagaactattgctgctatacgctaaataatcaat 1778
QY 1621 GGCTGTATGAGACAGCTATTTCTAACCGTCTGACACCACTTGTAGCAATGGCTTACGGT 1680
Db 1779 ggcctgtagcaagggcctatttcaacgctctgacacaccttgtagcaatgagccttacggt 1838
QY 1681 CGCTGCAAGATTACCGCTCTATGATGACTTACCTGCTGTAAGAGCAATCCAGAGA 1740
Db 1839 cgctgccaaggtttaccgctctatgtagaactaccttctgaaaggaagcaatccagagga 1898
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTATTTAAAAATGGTCTCG 1800
Db 1899 ttggaatataccagaggggctctacagaaatgagaaatctgtatttaaaatgctgcg 1958
QY 1801 TTCTACTGGAACACACTGCTGCTCCACAACAACCCCATCAACTAAGTTTAACTATC 1860
Db 1959 ttctaactgtagagctcactcctccacaacaccccccaactcgaatgaaagttcaaatcctatc 2018
QY 1861 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCAAGCAACAATATATAGTACGAC 1920
Db 2019 atcagatagttcaacttcacagttcagctcaaccatcccaagcaacaataatagtagac 2078
QY 1921 TACCAATCTTAACAATATATAGCAACAATCAAAATACAAACCCCTGATCAACAAAATCAGAA 1980
Db 2079 taccaatcctaacaataatagcaacaatcaaatcaaacccctgatacaacaataatcagaa 2138
QY 1981 TCTCAACGACGACACCA 1999
Db 2139 tcttcaaccagcacaccca 2157
```

## RESULT 5

AAS55688 standard; DNA: 2160 BP.

AAS55688:

13-FEB-2002 (first entry)

Streptococcus pneumoniae DNA for cellular proliferation protein #259.

Antisense: ds; prokaryotic cellular proliferation gene;

antibiotic; antibacterial; drug design.

Streptococcus pneumoniae.

MO200170955-A2.

27-SEP-2001.

XX

```
PF 21-MAR-2001; 2001MO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-267308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Heselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR P-PSDB; AAU37829.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27; Seq ID No 9325; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Streptococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2160 BP; 697 A; 522 C; 413 G; 528 T; 0 other;
```

Query Match 99.1%; Score 1981.4; DB 23; Length 2160;  
Best Local Similarity 99.4%; Pred. NO. 0;  
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
QY 1 TAAATCTACGACATAAAATCAATGCTGACTTGGTTCTGAACCGCGTCAA 60
Db 159 taaatctacgacaataaaatcaatcattgctgacttggttctgaaacgcgctcaa 218
QY 61 TGCCCAAGCTATATATTTCCACAGATTTGGTTAAGCAATGCTTTTATCGAAGACA 120
Db 219 tggccaagctaatatatttccacagatttggtttaagcaatgcttttatcgaaagaca 278
QY 121 TCGCTTCTTGACACAGAGGGGATTGATACCATCGTATCCTGGAGATTCTTGGCGAA 180
Db 279 tcgcttcttgacacagagggattgataccatcgatcctggagattcttggcgaa 338
QY 181 TCTGCAAAAGCAATTCCTCCAAAGGTGATCAACTCTCACCCCAAGTGTGATTAAGTTGAC 240
Db 339 tctgcaaaagcaattcctccaaggtgatcaactctcacccaagttgataagttgac 398
QY 241 TTACTTTTCAACTTCGACTTCGACAGAGACTATTTTCGTAAGGCTAGGAAAGCTTGTT 300
Db 399 ttacttttcaacttcgacttcgacagactatttctcgtaaggctgaagaaagcttgct 458
QY 301 AGGATTCACTTAGAACAAAAGCAACCAAGCAAGAATCTTGACTACTATTAATAATAA 360
Db 459 aggattcacttagaacaaaagcaacccaagcaagaatcttgactactataataataa 518
```

QY	361	GGCTACATGCTCTATGGGACCTATGGAAATGCACACAGCAGCTAAAACTACTATGGTAA	420
Db	519	ggctaccatgctctaaatgggaactatgaaatgcagacagcgtctaaaactactatgtaa	578
QY	421	AGACCTCAATATTTAAATTTAACTTCACTGATTTAGCCTTCTGGCTGGATGGCTCAGGCACC	480
Db	579	agacctcaataatttaagtttaacctcagtttagccttgctgctggaatgctctaagacc	638
QY	481	AAACCAATATGACCCCTTATTCATCCAGAAGCAGCCAGACCGCCGAACCTTGGTCTT	540
Db	639	aaaccaataatgacctctatcacatccagaagcagcccaagaaccgcgaacttgcctt	698
QY	541	ATCTCAATGAAAAATCAAGGCTACATCTCTGCGAAGCAGATGAGAAACAGTCATAC	600
Db	699	atctcaaatgaaaaatcaagctcaactctctgcgcgaaacgtaatgaaagcagtcataac	758
QY	601	ACCAATTTACTGATGACTTACAAAGTCTCAAAATGACAAAGTAAATTCCTGCTTACATGGA	660
Db	759	accaatttactgattgggctaccagaagttctcaaatccagaagtaattacctcttaactgga	818
QY	661	TAAATTAAGCAAGGATGATCATCATCAAGTTGAAGAAGAAACAGGCTATTAAGTACTCAC	720
Db	819	taattacctcaaggagatcatcaactcaagtatgaagaagaacagcgtataactactacac	878
QY	721	AACTGGATGATGCTCTACCAAAATGTAGACCAAGAAGCTCAAAACATGTGGGATAT	780
Db	879	aactggatgatagtctctacaaaatgtagaccacaagaagctcaaaaaactatgtyggatat	938
QY	781	TTACAAATACAGACGANTACGTTGCCCTTCCAGACGATGAATTCGACAGTGGCTTCAACAT	840
Db	939	ttacaatacagaagatacgttgccttaccagaacgaatgtcaagtcgtcttaccat	998
QY	841	TGTTATATTTTCTTAACGGTAAAGTCATTTGCCACCTGAGGACAGCCATAGTCAAGTAA	900
Db	999	tgttatagttcttcaacggtaaagtcattgcccagctagagaacgccaatcagtcaagtaa	1058
QY	901	TGTTTCTTGGAAATTTAAACAAGCAGTAGAAACAACCGGACTGGGATCAACTAGAA	960
Db	1059	tgttctcttcgaaatlaaccagaagtagaacaacacgcgacttggatccaactagaa	1118
QY	961	ACCGATGCACACATAGTCTCGGCTTGAGGTACGGGTCTAGANTCAACTGCTACTAT	1020
Db	1119	accgatacagaactatgctcctgccttggagtagtgcatacagatccactatgcactat	1178
QY	1021	CGTTACACANTAGGCGCTATTAATCTACCTGGGACAAATACTCTGTTTATTAATCGGATAG	1080
Db	1179	cgttacagaatgagccctataactaccccttggaacaataaccctgtttataactggatag	1238
QY	1081	GGGCTACTTTGGCAACATCACCTTGGCAATACGCCCTGCACAAATTCGGAAAAGTCCACG	1140
Db	1239	gggctacttggcaacatacactctgcaatacgccttgcacaacatcggaacgtcccaagc	1298
QY	1141	CGTGGAAACTCTAAACAAGTCGAGATCAACCGGCGCCACACTTCTCTAAATGGTGTAG	1200
Db	1299	cgtggaaactctaaacaagatcggaactccaacgcgcgaagaacttctctaaatggtctcgg	1358
QY	1201	AATGCATCCCAAGTATTCATCTCAATGCCATTTCAGATACACAACCGAATCAGA	1260
Db	1359	aatgcatacccaagatctactactcaaatgccaattccaatgaacaacaacggaatcaga	1418
QY	1261	CAAAAAATATGGACCAAGTGTGAAAAGATGGCGTCTAGCGTCCCTTGCCCAATGG	1320
Db	1419	caaaaaataatggacaagatgtaaaaaagatggcgctgcttaccgcgtcccttgcataatgg	1478
QY	1321	TGGAAGTACTATTAACCAATGTATATCCATAAAGTGTGTTTAAAGTATGGAGTGAATA	1380
Db	1479	tggagactactataaacacaatgatatacataaagtctgtttaaigtatggggtgtaaaa	1538
QY	1381	AGAAGTCTTAATGTGGAACTGTCGCCATGAAGAAGAACAGCAGCCTATATGATGACGA	1440
Db	1539	agaagttctctaaatgtcgaactcgtgcgaatgaagaagaacagcctatataatgtgcga	1598

QY	1441	CATGATGAAAAACGTCTTGACTTAATGAGCACTGGAGAAATGCTATCTTGTTGGCTCC	1
Db	1599	catgataaacaacagctcttgagttatatggaactggagcgaatgcttactcttgcttgagctccc	1
QY	1501	TCAGGCTGCTGTAACACAGAACCTCTTAACATATACAGACGAGAAATTTGAAAAACCATCAA	1
Db	1659	tcagagctgtaataaaccaggaaaccttaactatacagagagaaatctgnaaaccaatccaa	1
QY	1561	GACCTCTCATAATTTGTGTAGCACTGTGATGACTATTTTGGCTGTATACGCGTAATATTTCAT	1
Db	1719	gacctccaactattgtaacacctgataactattctgtcgctatacgcgtaataatccaat	1
QY	1621	GGCTGTATGAGACAGGCTATTTCTTAACCGTGTGACACCACTGTATAGCAATGGCCTTACGCT	1
Db	1779	ggctgtatlggacagagctattcttaaccgcttgtaaccacttgtagcaatlggcttaagct	1
QY	1681	CGCTGCACAAAGTTTACCGCTCTATGATGTGACCTTACCTGTGTGTAAGAAACATCCAGAGA	1
Db	1839	cgctgcacaaagtlttaaccgctctctatgtagcctaactgctctaaggaagaacatccagagga	1
QY	1741	TTTGAATATATACACAGAGGGGCTCTACAGAAATGGAGAATTTCGTATTTTAAAAATGGTGCTG	1
Db	1899	ttggaatatataccagaggggctcttaacgaatlgagagatctglatttaaaatggtgctg	1
QY	1801	TTTCTAGCTGGAACTCACTCTGCTCCACACACACCCCATCAACTGAAGTTCAAGCTATC	1
Db	1959	tctctagctggagagctcaactctctccacaacaccccatcaactgaagttcaagctcattc	2
QY	1861	ATTCGATATAGTTCAACTTCACAGCTGTGTGCTCAACCACTCCAGCAACAAATATATATAC	1
Db	2019	atcgaataagttcaactcaactcaagctcgaagcccaacccccaagacaataatagtagac	2
QY	1921	TACCAATCTCTAACAAATATATACGCAACATCAAAATCAACCCCTGATCAACAAAATCAGAA	1
Db	2079	taccaatccttaacaataataatagcaacaataatcaaatatacaacccctgatacaacaataatga	2
QY	1981	TCCTCAACCAACGACCAACCA 1999	
Db	2139	tcctcaaccagcacaacca 2157	
RESULT 6			
AAH02056			
AC	AAH02056	standard; DNA; 2160 BP.	
XX	AAH02056;		
DT	24-JUL-2001	(first entry)	
KW	Streptococcus pneumoniae	nucleotide sequence SEQ ID NO:2049.	
KW	Species specific; genus specific; family specific; probe; detection;		
KW	identification; algal; archaeal; bacterial; fungal; parasitic;		
KW	microorganism; diagnosis; translation elongation factor Tu; toxin;		
KW	translation elongation factor G; RecA recombinase; resistance;		
KW	catalytic subunit of proton-translocating ATPase; antimicrobial;		
KW	vaccine; primer; ds.		
OS	Streptococcus pneumoniae.		
XX	WO200123604-A2.		
XX	05-APR-2001.		
XX	28-SEP-2000; 2000WO-CA01150.		
XX	28-SEP-1999; 99CA-2283458.		
XX	19-MAY-2000; 2000CA-2307010.		
XX	(INFE-) INFECTIO DIAGNOSTIC (IDI) INC.		
XX	Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;		
XX	Picard FJ, Roy PH.		
XX	1		



QY 1561 GACCTCAATTGTTGACGACCTGATGAACTATTGCTGCTATACGGCTAAATATTCAT 1620  
DB 1719 gaacacgtgcatgtagctccagaaatggttctggtatcctgtagatctctat 1778  
QY 1621 GGCTGATGAGCAGGCTATCTTAACCGCTGACACCACTGTGAGCAATGGCTTACGCT 1680  
DB 1779 ggcctgcatgagcaggtatcgaatcgttcaactccatggttggagatggttccagt 1838  
QY 1681 CGCTGCCAAGTTTACCGCTCTATGACCTGCTGTAAGAAAGCAATGTCAGAGA 1740  
DB 1839 tgcagctaaagttatcgctccatgatacgtatctatcagaagatactatccagaaga 1898  
QY 1741 TTGGATATATACGAGGCGCTTACAGAAATGGAATTCGATTTTAAATAATGGTCTCG 1800  
DB 1899 ctgagagatgctcagcaggaacttccagaaatggagaatcgtatttaaaatggtctg 1958  
QY 1801 TTCTGATGGAATCAACGCTCTCACAACACCCCATCATCAAGAAAGTTCAAGCTCATC 1860  
DB 1959 tctcaagtggaactccgctccacaacacccccaactgaagttcaagctcatc 2018  
QY 1861 ATCAGATAGTTCAACTTTCACAGTCTAGCTCAGACCTCCAGCAAAATATGATGAC 1920  
DB 2019 atcagtagtcaactcaactcaagctcgtatcaaccactccagcaataatagtctagc 2078  
QY 1921 TTACCAATCTTACATATATGACCAATCAATCAATCAATCAATCAATCAATCAATCA 1980  
DB 2079 taccatcccaacaataatagcaacaatcaatacaccctcgatcaacaataatcaga 2138  
QY 1981 TTCTCAACGACGACCAACCA 1999  
DB 2139 tcccaacagcagacaaca 2157

RESULT 7

AAH01178  
ID AAH01178 standard; DNA; 2160 BP.  
AC AAH01178;  
XX  
XX  
XX 24-JUL-2001 (first entry)  
XX  
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1169.  
XX  
XX Species specific; genus specific; family specific; probe; detection;  
XX identification; algal; archaeal; bacterial; fungal; parasitica;  
XX microorganism; diagnosis; translation elongation factor Tu; toxin;  
XX translation elongation factor G; RecA recombinase; resistance;  
XX catalytic subunit of proton-translocating ATPase; antimicrobial;  
XX vaccine; primer; ds.  
XX  
OS Streptococcus pneumoniae.  
XX  
XX  
XX WO200123604-A2.  
XX  
XX 05-APR-2001.  
XX  
XX 28-SEP-2000; 2000WO-CA01150.  
XX  
XX 28-SEP-1999; 99CA-2283458.  
XX 19-MAY-2000; 2000CA-2307010.  
XX  
XX (INFE-) INFECTION DIAGNOSTIC (IDI) INC.  
XX  
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;  
XX Picard FJ, Roy PH;  
XX  
XX WPI: 2001-245006/25.  
XX  
XX Nucleic acid sequences are used to generate universal probes and  
XX primers which can be used to identify and detect the presence of algal,  
XX archaeal, bacterial, fungal and parasitica species in a test sample -  
XX  
XX Disclosure: Page 1044-1045; 1580bp; English.

XX  
CC The present invention describes a method for generating a repository of  
CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes  
CC and/or primers are derived. The method comprises amplifying the nucleic  
CC acids of determined algal, archaeal, bacterial, fungal and parasitica  
CC species with a combination of defined primer pairs. The method can be  
CC used for producing probes and/or primers for detecting one or more  
CC related microorganisms e.g. algae, archaea, bacteria, fungi and  
CC parasites, for universal detection and for specific and ubiquitous  
CC detection and identification of an algal, archaeal, bacterial, fungal  
CC and parasitica species, genus, family and group. A nucleic acid (I)  
CC obtained using the method of the invention can be used for the universal  
CC detection of any bacterium, fungus or parasite in a sample and for the  
CC detection of at least one antimicrobial agent resistance gene or at  
CC least one toxin gene. hexA nucleic acids are used for the specific and  
CC ubiquitous detection and for identification of Streptococcus pneumoniae.  
CC (I) can be used to design a therapeutic agent which is effective against  
CC microorganisms. Microbial species or genus or family or phylum or group  
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,  
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,  
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,  
CC Neisseria gonorrhoeae and staphylococcus sp.. Using DNA based tests  
CC provides faster results than substrate specificity tests as results can  
CC be determined in an hour and improved accuracy is also achieved.  
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes  
CC which are given in the exemplification of the present invention.  
XX  
SQ Sequence 2160 BP; 698 A; 494 C; 403 G; 565 T; 0 other;

Query Match 82.2%; Score 1642.2; DB 22; Length 2160;  
Best Local Similarity 88.8%; Pred. No. 0;  
Matches 1776; Conservative 0; Mismatches 223; Indels 0; Gaps 0;

QY 1 TAAATCTGACGACAAATTAATCACTGCTGACTTGGCTTGAAGCGCGCTGCA 60  
DB 159 taaatctgacgacaaatlaaataactctgctgacttggcttgaagcgcgctgca 218  
QY 61 TGCCCAAGCTAATGATATTCGCCAGATTTGTTAAGGCAATGTTTATGCAACCA 120  
DB 219 tgcccaagctaatgatatccccaagattggttaagcaatcgtttcatcgaagacca 278  
QY 121 TCGCTTCTTGACACAGGGGGATTGATACATCCGTATCCCTGGAGCTTCTTGGCA 180  
DB 279 tcgcttcttgacacaggggattgatcatccgtatccctggagcttcttggcaa 338  
QY 181 TCTGCAAGCAATTCCTCTCAAGGTGATCACTCTCACCACCAATGTTAAGTTGAC 240  
DB 339 tctgcaagcaatctcctctcaaggtgatcatctcaccaccaagttgattaagttgac 398  
QY 241 TTACTTTTCAACTTGCACTTCCGACGACTATTTCTCGAAGGCTCAGAACTTGCTT 300  
DB 399 ttacttttcaacttgcacttccgacgactatttctcgaaaggctcagaaacttgctt 458  
QY 301 AGCGATTGAGTGAACAAAGCAACCAAGCAAGAAATCTTGACCTACTATTAATAA 360  
DB 459 agcgattgagtgaaacaaagcaaccaaagcaagaaatcttgactactataataaa 518  
QY 361 GGCTTACATGTCTAATGGCACTATGGAATGCAGACAGCAAGCTCAAAATCTATGTTAA 420  
DB 519 ggccttacatgtctaatggcactatggaaatgcagacagcaagctcaaaatctatgttaa 578  
QY 421 AGACCTCAATATTTAAGTTTACCTAGTTAGCTTGCGTGGGGAATGCTCAGGCAC 480  
DB 579 agacctcaatatttaagtttacctagtttagcttgcggtgggaatgctcaggcac 638  
QY 481 AAACCAATATGACCCCTATTTACATCAGAGCAGCCCAAGACCGCGCAACTTGGCTT 540  
DB 639 aaaccaatatgacccctattttacatcagagcagcccaagacgcgcaacttggctt 698  
QY 541 ATCTGAATGAATAATCAAGGCTACTCTGCTGACAGTATGAGAAAGCACTAATAC 600  
DB 699 atctgaatgaataatcaaggctactctgctgacaggtatgagaaagcactaataac 758

QY 601 ACCATTACGATGAGTACAAAGTGCATCAATGCAAGCAAGTAATTACCTGCTACATGGA 660  
DB 759 accaattacgatacgatacaaaagtcctcaaaatcaagcaagtaataccctgcttacaatgga 818  
QY 661 TAATTACCTCAAGGAAGTATCATCAATCAAGTGAAGAAGCAAGGCTATATCACTACAC 720  
DB 819 taattacctaagaaggtacatacaatgaagtaagaacaagaactgctataacttctaac 878  
QY 721 AACTGGAGTATGCTCTACCAAAATGTAGCCAGAAAGCTCAAAAACATCTGTGGATAT 780  
DB 879 tactggatgatagtttacacaatgtagacaagaagctcaaaaactctggyggtat 938  
QY 781 TTTCATACAGAGAGATACCTTGCTTCCATCCAGAGATGATTTGCAAGTCCCTCTACCAT 840  
DB 939 ctacaactccgaatacaatgctcttaccctgacgaatgattgcaagctgcatcacaagt 998  
QY 841 TGTGTATGTTTCTAAGCGTAAAGTCAATGCTCCAGCTAGAGACAGCCATCAGTCAAGTAA 900  
DB 999 cgtagatglttcaaatgataaagtcatacgccaacttgagagctcgltcaacaagaagtaa 1058  
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGATCAATATGAA 960  
DB 1059 cgttcaattgtagacaacccaagctgtagaacaacatcgtagctggygttctgctatgaa 1118  
QY 961 ACCGATCAGAGACTATGCTCTGCTTGGATAGGCTGTCTACGATTTCACTGCTACTAT 1020  
DB 1119 accaatcaacgattatgacacctgacataagaaatagatgatttaccatccatgcaaat 1178  
QY 1021 CGTTACGATGAGCCCTATATACCTACCTGGAGCAAAATACTCTGTTATATACAGGATAG 1080  
DB 1179 ggttaattgataatccttatactatactccggaagaacacccgtcttacaactggygtag 1238  
QY 1081 GGGCTACTTGGCAACATCACTTGCATACGCTTGCACAAATCGGCAAAAGCTCCAGC 1140  
DB 1239 agaatattcgtaataatactctgcaataatgctcttcaacaatacgaatgtcacagc 1298  
QY 1141 CGTGAACACTCTAACAAGGTCGAGCTCAACCCGCCAACAACCTTCTTAATAGTGTCTAG 1200  
DB 1299 cgttgagacttgaataaagtcgctgacataagaaacacttcccttaagctctg 1358  
QY 1201 AATGACTACCCAGTATTTACTACTCAATGCAATGCAATTTCAAGTAACCAACCGAATCAGA 1260  
DB 1359 taccgactatccaagatgattatgcaaacgcatcttaagtaatacaacagaattcaa 1418  
QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAAAGATGCTGCTTACGCTGCTTTGCAAAATGG 1320  
DB 1419 taacaatacagagcaagtagtgaaaaaatgctgcttactgcttgccttgcaaatg 1478  
QY 1321 TGGAACTTACTATAAACCATGTATATCCATAAAGTGTCTTATAGTATGGAGTGAATA 1380  
DB 1479 tggcaacttactataaaccatgataccaataaagctgcttactagtagaagtaaa 1538  
QY 1381 AGAGTCTCTAATGTCGGAAGCTGTCGCATGAAGAAAGCAAGCCATATATGATGACGA 1440  
DB 1539 agagttctctaaatgctggaactcgtgcatagaagaagacagcctataatgtagcga 1598  
QY 1441 CATGATGAAAAAGCTTGAATTATGGAAGTGAAGAAATGCTATCTTGGCTGGCTCC 1500  
DB 1599 catgatagaagaagcttgccttactggaactggtgagcctatcttcccttggtcc 1658  
QY 1501 TCGAGCTGGTAAAAAGCAAGCTCTAATCTATACAGAGAGAAATTTGAAACCATCA 1560  
DB 1659 tcaagctgataaaacaggaaccttactatacagatgagaagttgaaaaacacatacaa 1718  
QY 1561 GACCTCTCAATTTGTAGCAGCTGATGAATTAATTTGCTGTATACGAGGTAAATTTCAAT 1620  
DB 1719 gaacctggtctatgtagctccagataagaatgltgtgtgtataactcgtataagtaattcat 1778  
QY 1621 GGGCTGATGAGAGGCTATTTCTAACCCTGTGACACCACTTGTAGGCAAGTGGCTTACG 1680  
DB 1779 ggtctgataggaacaggtattcgtatcgatcgtttaactcctatcgltgagatagtttccag 1838

QY 1681 CGGTCCCAAGTTTACCGCTCTATGATGATACCTGCTGTGAAGAGCAATCCAGAGA 1740  
DB 1839 tgcagctaaagttatgctcacaatgataacgatactctacagaagatcctacagagaa 1898  
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAAAATGGTGTCTG 1800  
DB 1899 ctgagcagatgcagagacttttcagaacaggggaattgtatttcaaaaatgagctcg 1958  
QY 1801 TTCTACGTGCACTACCTCTCTCCACAGAACCCCACTCACTGAAAGTTCAAGTCAATC 1860  
DB 1959 cccaataatgactgaacccctactcacaataatcctcacaagctgaaagttcaagctcatc 2018  
QY 1861 ATCAGATAGTTTCACTTACAGTCTACGCTTACCTACACACCTCAAGCAATATATAGTAC 1920  
DB 2019 atcagataatgctcactcacaatgctagctcacaaccccaagcaacaataatagtagac 2078  
QY 1921 TACCAATCTTACAAATTAATAGCAACAAATCAATCAATCAACCCCTGATCAACAAATCGAA 1980  
DB 2079 tacaatatcctaacaataatacgacaacatacaatacaacccctgatacaacaataatga 2138  
QY 1981 TCCTCAACGACACACCA 1999  
DB 2139 tcctcaacgacacacca 2157

## RESULT 8

AAH02054 standard; DNA: 2160 BP.

AAH02054:

24-JUL-2001 (first entry)

Streptococcus pneumoniae nucleotide sequence SEQ ID NO: 2047.

Species specific; genus specific; family specific; probe: detection;

identification; algal; archaeal; bacterial; fungal; parasitic;

microorganism; diagnosis; translation elongation factor Tu; toxin;

translation elongation factor G; RecA recombinase; resistance;

catalytic subunit of proton-translocating ATPase; antimicrobial;

vaccine; primer; ds.

Streptococcus pneumoniae.

WO200123604-A2.

05-APR-2001.

28-SEP-2000: 2000MO-CA01150.

28-SEP-1999: 99CA-2283458.

PR 19-MAY-2000: 2000CA-2307010.

(INFE-) INFECTIO DIAGNOSTIC (IDI) INC.

Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;

Picard FJ, Roy PH;

WPI: 2001-245006/25.

Nucleic acid sequences are used to generate universal probes and

primers which can be used to identify and detect the presence of algal,

archaeal, bacterial, fungal and parasitic species in a test sample -

Disclosure: Page 1462-1463: 1580bp: English.

The present invention describes a method for generating a repository of

nucleic acids of tuf, fus, atpD and/or recA genes from which probes

and/or primers are derived. The method comprises amplifying the nucleic

acids of determined algal, archaeal, bacterial, fungal and parasitic

species with a combination of defined primer pairs. The method can be

used for producing probes and/or primers for detecting one or more

related microorganisms e.g. algae, archaea, bacteria, fungi and







QY 1801 TTCTAGCTGGAACTGACCTGCTCCACACACACCCCATCATCTGAAAGTTCAAGCTCATC 1860  
 DB 1959 tttctacgtgagagctaccgcgtcccaacaacccccatcaactgaagtcaagctcacc 2018  
 QY 1861 ATCAATAGTCTTAACCTTCACTGCTCAACCACTCCACACAAATATATGATGAC 1920  
 DB 2019 atcagatagttcaactcaagctcaactcaacacactccaagcacaataatagtagcagc 2078  
 QY 1921 TACCAATCTTACATATATATGCAATCAATCAATCAATCAATCAATCAATCAATCA 1980  
 DB 2079 taccatctcacaataaatacgaacaatacaatacaacccctgatacaacaatacga 2138  
 QY 1981 TCCCTACACGACCAACCA 1999  
 DB 2139 tctctacacagcacaacca 2157

RESULT 9  
 AAH02147  
 ID AAH02147 standard; DNA; 2157 BP.  
 AC AAH02147;  
 DT 24-JUL-2001 (first entry)  
 DE Streptococcus pneumoniae nucleotide sequence SRO ID NO:2140.

KM Species specific; genus specific; family specific; probe; detection;  
 KM Identification; algal; archaeal; bacterial; fungal; parasitical;  
 KM microorganism; diagnosis; translation elongation factor Tu; toxin;  
 KM translation elongation factor G; RecA recombinase; resistance;  
 KM catalytic subunit of proton-translocating ATPase; antimicrobial;  
 KM vaccine; primer; ds.  
 OS Streptococcus pneumoniae.  
 PN MO200123604-A2.  
 PD 05-APR-2001.  
 PE 28-SEP-2000; 2000WO-CA01150.  
 PR 28-SEP-1999; 99CA-2283458.  
 PR 19-MAY-2000; 2000CA-2307010.  
 PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.  
 PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;  
 PI Picard FJ, Roy PH;  
 DR WPI: 2001-245006/25.  
 PT Nucleic acid sequences are used to generate universal probes and  
 PT primers which can be used to identify and detect the presence of algal,  
 PT archaeal, bacterial, fungal and parasitical species in a test sample -  
 PS Disclosure: Page 1508-1509; 1580pp; English.

CC (1) can be used to design a therapeutic agent which is effective against  
 CC microorganisms. Microbial species or genus or family or phylum or group  
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,  
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,  
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,  
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests  
 CC provides faster results than substrate specificity tests as results can  
 CC be determined in an hour and improved accuracy is also achieved.  
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes  
 CC which are given in the exemplification of the present invention.  
 XX  
 SO Sequence 2157 BP; 696 A; 494 C; 405 G; 562 T; 0 other;

Query Match 80.4%; Score 1607; DB 22; Length 2157;  
 Best Local Similarity 87.7%; Pred. No. 0;  
 Matches 1754; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATTAATCAACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60  
 DB 159 taaatctacgacaataaatacaactcatctgctgagctgggtctgcaagcgcgctcaa 218  
 QY 61 TGCCCAAGCTTAATGATATTCACAGATTTGGTTAAGGCAATCGTTTATCGAAGACCA 120  
 DB 219 tgcccaagctaatgatattccacagatttggttaagcaatcgtttctatcgaagacca 278  
 QY 121 TGGCTTCTTCGACACACAGGGGATGATACATCCGTATCTGCGGACCTTTCTTGCGCAA 180  
 DB 279 tggcttcttcgacacacaggggattgattccatccgtatccgggagcttcttgcgcaa 338  
 QY 181 TCGCAAAACCAATTCCTCCCAAGGTGATCAACTCTCACCAACAGTTTATTAAGTTGAC 240  
 DB 339 tctgcaaaatattccctcccaaggtggaaccaactccaccaagtgattaaagtgcac 398  
 QY 241 TTACTTTTCAACTGCACTTCGACCACTATTTCTCTAAGGCTCAGAAAGCTTGCTT 300  
 DB 399 ttactttcaactcgacttcgaccagactattctctgtaagctcaggaagcttggtt 458  
 QY 301 AGCGATTGCTTGAAGCAAAAGCAACGACGAAGAAATCTTGACTACTATATTAATTA 360  
 DB 459 agcgatttcgcttgaagcaaaagcaacgaaagaaatcttgactactatataataaa 518  
 QY 361 GGGCTACATGCTTAATGGAAGTATGACAGTGCAGACAGCGTCAAAATCTACTATGTTAA 420  
 DB 519 gggctacatgcttaatggaaagtatgacagtgcagacagcgtaaaatctactatgtaa 578  
 QY 421 AGACCTCAATTAATTAAGTTTACTCTAGTTAGCTTCTGCGGTGATGCTTCAGGACCC 480  
 DB 579 agacctcaaatatttaagtttaactctagtttagcttgcgtgcgtggaatgctccagggcac 638  
 QY 481 AAACCAATATGACCCCTATTCACATCCAGACAGCCCAAGACCGCGAACTTGCTTT 540  
 DB 639 aaaccaatattgacccctatttcacatccagaagcagccagcggaacttgctt 698  
 QY 541 ACTGTAAATGAATAATCAAGGCTACATCTCTGCTGAACGATGTAAGAAAGCTCAATAC 600  
 DB 699 atctgaaatgaataatcaaggttcaatctctgctgacagatlgaaagcgactcaatc 758  
 QY 601 ACCAATTAATGATGACCTCAAAAGTCTCAATCAAGCAAGTAAATTAATCCCTGCTTAC 660  
 DB 759 accaatttaattgactcaaaagctcaatcagaagaagttaattaccctgcttaactgga 818  
 QY 661 TAATTAAGTCAAGGATGATCAATCAAGTGAAGAAAGCGTATTAACCTACTAC 720  
 DB 819 taattaaagtcgaaggtatcaatcaatgaagaagaactgacttaaaccttctaac 878  
 QY 721 AACTGGATGATGATGCTTACAAATGTAGCAAGCAAGCTCAAAACATCTGGGATAT 780  
 DB 879 taactggatgattgcttacaacaatgtagaccagaagctcaaaaacatctgtaggatac 938  
 QY 781 TTACATACAGAGATATAGCTTATCCAGACGATGATGATGATGATGATGATGATGAT 840  
 DB 939 ctacaactcgttcaataagctcttcttacccttgacgtatgttgcaagtcgacttaaggt 998

QY	841	TCGTTGATGTTTCTAAGCTTAAAGTCATTTGGCCAGCTAGGAGGACGCCATTCAGTCAAGTAA	900
Dp	999	cgtagagcttccaatggttaaaagctatcgccaaactctgtagctgtccacgaagcaaa	1058
QY	901	TGTTTTCCTGGGAATTAACCAAGCAGTAGAAGAACCAACCCGACTGGGATCAACTATGAA	960
Dp	1059	cgcttcatttggtaaccaacgaagctgtggaanaaccaatcgtagctgggttcgtcatgaa	1118
QY	961	ACCGATCACAGACTATGCTCTCTGCTTGGAGTACGGTGTCTACGATTCACATGCTACTAT	1020
Dp	1119	accataaccgatatgacacctgcacatgaaatcggtttctatgattccactgcgaactat	1178
QY	1021	CGTTACAGATGAGCCCTTAATCTACCTGGGAGCAAAATATCTCTGTTTAACTGGGATAG	1080
Dp	1179	ggttaatgatatcttcctataactatcccggaataaagcacactgtcttaactcgttag	1238
QY	1081	GGGCTACTTTGGCACAATCACTTTCGAATATACGCCCTGCACAAATCGGAACGTCCACG	1140
Dp	1239	agcaatctcgttaatatctctctgcaatatgctcttcacaatacgaatgltcacagc	1298
QY	1141	CGTGGAAACTCTAACAAGGTGGAGCTCAACCGCGCAAGACTTTTCTCAATAGGCTTGG	1200
Dp	1299	cgtagaactcttgaaataagctcggtccagaataagctaaaccttccttaacgtctcgg	1358
QY	1201	AATCGACTACCCAAATTTTACTACTCAATATGCGATTTTCAGTAAACACAGCATCAGA	1260
Dp	1359	tatcgactatccaagcatgcatatgcaaacgcattcgaatgaataacagaagatctaa	1418
QY	1261	CAAAAATATGAGAGCAAGTGTGMAAAGATGGCTGTGTTCGCTTGGCTTTGGCAATGG	1320
Dp	1419	taaacataacggagcaagtagtgaaaaaagtgctgtccttatgctgccttgcacaag	1478
QY	1321	TGGAACCTACTATTAACCAATGTATATTCATAAAGTCTGCTTATAGTAGAGAGTAA	1380
Dp	1479	tggaaattcaccaacaaacccatgatatcaataaagtcgtcttcagtgagcgtgagtaaaa	1538
QY	1381	AGAATTCTTAATGTGGAACTGTGCCATGAAGAAGACAGACTTATGATGACCGA	1440
Dp	1539	agaaattctcagatgtagtagcacgagctatgaaagaacaaactgtctacatgatlbaacga	1598
QY	1441	CATCATGAAAAAGCTGTGCAATTTATGAACTGGACAAATGCCPTTCTTGCTGGCTCC	1500
Dp	1599	aatgatagaaaactcgtctcgtgacaagaaactcgtctcgttgaagcctatctcccatggttcg	1658
QY	1501	TCAGGCTGTGTAACAGGACACTCTTACTATACAGACGAGGAATTTGAAAACCATCA	1560
Dp	1659	gcaagctcgttaagaacaggtactcttcaactacaagaatgaaatctgaaaaacaacataa	1718
QY	1561	GACCTGCATTTGTAGCACCCTGATGACATTTTGTGTGCTTACGGCTAAATATTCAAT	1620
Dp	1719	gaaacaacggtcaatgtagctcccaagatgaatgltgttgttatactcgttaagatattcat	1778
QY	1621	GGCTGTATGACAGGCTATTTCTAACCGTGTGACACACTTTGAGCAATGGCCTTACGGT	1680
Dp	1779	ggtctgtagtgaacaggtattctcgatcggtttaactcctatcgttggagtggtttccatg	1838
QY	1681	CGCTGCCAAAGTTTACCGCTCTATGATGACCTTACTGTGTGAAGACAGCAATCCAGAAGA	1740
Dp	1839	tgcagctcaaaagttatcgtctcaatgataaagctatcatbaagaatactcatcacaaga	1898
QY	1741	TTGGAATATACCAAGAGGGCTCTACAGAAATGGAGAATTCGATTTAAATAATGGGCTCG	1800
Dp	1899	cttagacgaatgcacgaaggaactttccagaaacgggaatttgtattccaataatlgagctcg	1958
QY	1801	TTCTAGCTGGAACCTCACCTGCTTCACAGAACCCCCATCACTGAAGTTCAAGCTATC	1860
Dp	1959	tctcaagctggagctcacacctctccacaacaaccccacatcaatgaaatcaagctatc	2018
QY	1861	ATTCGATATGTTCAACTTCACAGCTGTAGCTCACACCACTCCAAACAATAATATAGACAC	1920
Dp	2019	atcagaaatagttcaactcaacagtctagcttcaacaaactccaagcaaaaaatagtagcac	2078

OY	1921	TACGCAATCCTAACAAATAAATAGCGACAATTCGAATACCCTGGATCAACAATAATGCAGA	1980
XZ			
DB	2079	Taccgactgtgaataacaaataaacgcacaatcaaatcacaacctccgatacaacaattcacaga	2138
OY	1981	TTCCTCACACGACGACACACCA 1999	
XZ			
DB	2139	tctctaaccagcacataacca 2157	
 RESULT 10 AAZ35952 ID AAZ35952 standard; DNA; 1260 BP. XX XX AAZ35952; AC DT 07-FEB-2000 (first entry) DE Streptococcus pneumoniae pbp1A TER isolate n) nucleotide sequence.			
KM	Streptococcus pneumoniae: penicillin binding protein; pbp2B; pbp1A;		
KW	transpeptidase encoding region; TER; antibiotic resistance; diagnosis;		
KM	detection; identification; pneumococcal meningitis; ss.		
OS	Streptococcus pneumoniae.		
PN	ZA9807024-A.		
PD	28-APR-1999.		
PX	05-AUG-1998; 98ZA-0007024.		
PF	01-AUG-1997; 97ZA-0006886.		
PR	(SAME-) SOUTH AFRICAN INST MEDICAL RES.		
PA	(UYWI-) UNIV WITWATERSRAND.		
PA	(MED-I-) MEDICAL RES COUNCIL.		
PI	Klugman KP, Smith AM, Du Plessis M;		
DR	WPI; 1999-601770/51.		
PT	Polymerase chain reaction assays for detecting Streptococcus pneumonia		
PP	useful for the diagnosis of pneumococcal meningitis -		
PS	Claim 11; Fig 4; 63pp; English.		
CC	A polymerase chain reaction (PCR) assays have been developed for		
CC	detecting an antibiotic resistant strain of Streptococcus pneumoniae		
CC	using primers based on the penicillin binding Protein 2B (pbp2B) gene		
CC	and the pbp1A gene. The products and methods can be used for detecting		
CC	S. pneumoniae, particularly antibiotic-resistant strains. They can be		
CC	used for simultaneously diagnosing pneumococcal meningitis and		
CC	identifying any antibiotic-resistant S. pneumoniae strains in a sample.		
CC	The methods can be used for detecting S. pneumoniae strains resistant		
CC	to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.		
CC	The assays can be adapted to detect other pathogens causing meningitis.		
CC	The assays can be used to detect an antibiotic resistant strain of		
CC	S. pneumoniae with a minimum inhibitory concentration (MIC) of		
CC	0.25-1 micro g/ml where the PCR products detected are a 1043 bp and		
CC	a 224 bp product. The present sequence represents a Streptococcus		
CC	pneumoniae pBP1A transpeptidase encoding region (TER) isolate nucleotide		
CC	sequence from the present invention.		
SO	Sequence 1260 BP; 407 A; 292 G; 259 G; 302 T; 0 other;		
 Query Match      62.6%; Score 1252; DB 20; Length 1260; Best Local Similarity    99.6%; Pred. No. 0; Matches 1255; Conservative    0; Mismatches    5; Indels    0; Gaps    0			
OY	575	GAAACGTATGAAAAGCAGACTCAATATACCAATTACTGATGAGTAAAGATCAATCA 634	
XZ			
DB	1	gaaacagtbtggaaaaggcaatcaatacaacttcgatgtagtggaatcaaaagtccaatca 60	

OY	635	GCAGTAATTAACCTGCTTACATGGAATTAATTAACCTCAGAGAACTCATCAATCAAGTTGAA	694
Db	61	gcaagtaattacccttgcttaccatgataagtaattaccctcaagaagtcataatcaagtgtgaa	120
OY	695	GAAAGAAAGGCTTAATCACTACACAACTGGGATGGATGATGTCATACCAAAATGTAGACCA	754
Db	121	gaagaaacaaggctataactactacacaaactggagtagtgcctacacaaatgtagccaa	180
OY	755	GAACTCAAAAACATCATGTGGGATTAATTAATAATACAGACAAATACCTTGCTATCCAGAC	814
Db	181	gaagctcaaaaactcctcgttgggataattacaataacagaaagataagtttccatccagac	240
OY	815	GATGAATTCGAAGTCGCTTCTACCACTTTGATGTTTCTTAACGGTAAGTCATTTGGCCAG	874
Db	241	gatgatattgcaagtcgcgtcttaccattgttgatgtgttctaaggtaaagtcattgtccag	300
OY	875	CTAGGAGACCCCATCAAGTAAATGTTTCTTGGAATTAACCAAGCAGTAGAACA	934
Db	301	ctaggagacgcacatcagtaagtaagtcttcttccttcgaaatlaaccaagcagtagaaca	360
OY	935	AACGGCGATCGGGATCAACATTAATTAACCATACAGACTATGCTCGTCGTTGGAGTAC	994
Db	361	aaccggaacttggggataactaactagaacacgatacaagactatgtcttcgccttagatgc	420
OY	995	GGTGCTTACGATTCACCTGCTACTATTCGTTACAGATGAGCCCTTAATCACTCCGTGGACA	1055
Db	421	ggtgtctacgattctcaactgtcactatcgtltcacgataagcctataactacccttggaca	480
OY	1055	AATACCTCCTGTTTATACCTGGGATAGGGGCTACTTTGGCAACATCACTTTGCATACGC	1111
Db	481	aatacctctgttlaataactcgtgagtaggggtactcttggcaacatacacttgaataatgcg	540
OY	1115	CTGGAACAATGCGCAAAAGTCGCCAGCCGTGGAAACTTAACAAGGTGCGAGCTCAACGCC	1175
Db	541	cttgcacaacatcgcgaaacgctccagccgttggaaacctctaacaagttcggactcaacgcg	600
OY	1175	GCCAAAGCTTTCCTTAATGCTTAGAANTGACATACCCAGTAATTCATCACTCAATATGCC	1233
Db	601	gcacaagcttcttctaataatgttcttaggaatcgactaccacaagtatctcaactccaatgic	660
OY	1235	ATTTCAGTAAACACACCCGATCAGACAAAATAATGAGACAAGTAGTGAANAAGATGCT	1299
Db	661	attccaagtaacaacaacgcgaatcagacaaaaatatvgagcaagtagtgaagaatgagct	720
OY	1295	GCTGCTTAACGCTGGCTTTGCAAAATGCTGGAACCTTACTATTAACCATGTATATCCATAA	1355
Db	721	gtcttcttaagctgcgcttcttgaanaatggtggaacttactataaacaatgtatataccaata	780
OY	1355	GTCGCTTTAGTGAATGGAGATGGAATAAGATTTCTTAATGTCGAGACTGTCGCATGAAG	1414
Db	781	gtgcctcttagttagtggagtgtaaaaaagttctctcaatgttcggaactcgttgcataagag	840
OY	1415	GAAACGACAGCCTTAATGATGACCGACATGATGAATAAACAAGCTTTGACTTAGAAGATGGA	1477
Db	841	gaacagcagcctataatgataagcagacatgatagaacaacagttcttggattatgaaactgga	900
OY	1475	CGAATGCTATCTTTGTTGGCTCCCTCAGGCGGTGTAAACAAGAACGCTTCAACTATACA	1533
Db	901	cgaatgtcctactcttctgttgcctcccaagtcgtgtaaaaaacggaaccccttaactataca	960
OY	1535	GACAGAGAAATTGAAAACCATCAAGACCTCTCAATTTGTATACACCTGATGAACATATT	1599
Db	961	gagcagagaaattgaaaacccaalcaagaacctctccaattttagcacccttgaagcaactattt	1020
OY	1595	GCTGGCTATACGCGTAATAATTCATATGGCTGTATGACACAGCCTATTTCACGCTGAC	1655
Db	1021	gtctgctatacgcgtaaatatcaaatgtcgtatgtgaacagcgtattcttaacgcgttgaca	1080
OY	1655	CCACTTTGAGGCATGGCCCTTAGCGTGGCTGCCAAGTTTACCGCTCTATGATGACTAC	1711
Db	1081	ccacttttagagcaatggccttcttagtgctgcgcgaagaatttaacgcctctatgtatgactac	1144

OY	1715	CCTGCTGAAGAGCAATCCACAAAGATTGGAATTATTCACGAGGCGCTGTACAGAATAATGCA	1774
Db	1141	ctgtctgaagagaacatccagaggattggaatataccagagggcctctacagaata199a	1200
OY	1775	GAAATTCGATTTTAAAAATGCTGCTGTTCTACGTGGAACTCACGCCTCCTCAACAACC	1834
Db	1201	gaatcgtaattaaataatgctgcgtctctactacgltggaactcacctgcttcacaacaacc	1260
 RESULT 11 AA235950			
ID	AA235950	standard; DNA; 1260 BP.	
XX	AA235950;		
XX	07-FEB-2000	(first entry)	
DE	Streptococcus pneumoniae pbp1A TER isolate 1) nucleotide sequence.		
KW	Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;		
KM	transpeptidase encoding region; TER; antibiotic resistance; diagnosis;		
KW	detection; identification; pneumococcal meningitis; ss.		
XX			
OS	Streptococcus pneumoniae.		
XX			
PN	ZA9807024-A.		
SC	28-APR-1999.		
PD			
XX			
PF	05-AUG-1998; 982A-0007024.		
PR	01-AUG-1997; 97ZA-0006886.		
PA	(SAME-) SOUTH AFRICAN INST MEDICAL RES.		
PA	(UYWI-) UNITV WITWATERSRAND.		
PA	(MED1-) MEDICAL RES COUNCIL.		
XX			
PI	Klugman KP, Smith AM, Du Plessis M;		
DR	WPL; 1999-601770/51.		
XX			
PT	Polymerase chain reaction assays for detecting Streptococcus pneumonia		
PR	useful for the diagnosis of pneumococcal meningitis -		
XX			
PS	Claim 11; Fig 4; 63pp; English.		
XX			
CC	A polymerase chain reaction (PCR) assays have been developed for		
CC	detecting an antibiotic resistant strain of Streptococcus pneumoniae		
CC	using primers based on the penicillin binding protein 2B (pbp2B) gene		
CC	and the pbp1A gene. The products and methods can be used for detecting		
CC	S. pneumoniae, particularly antibiotic-resistant strains. They can be		
CC	used for simultaneously diagnosing pneumococcal meningitis and		
CC	identifying any antibiotic-resistant S. pneumoniae strains in a sample.		
CC	The methods can be used for detecting S. pneumoniae strains resistant		
CC	to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.		
CC	The assays can be adapted to detect other pathogens causing meningitis.		
CC	The assays can be used to detect an antibiotic resistant strain of		
CC	S. pneumoniae with a minimum inhibitory concentration (MIC) of		
CC	0.25-1 micro g/ml where the PCR products detected are a 1043 bp and		
CC	a 224 bp product. The present sequence represents a Streptococcus		
CC	pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide		
CC	sequence from the present invention.		
XX			
XX			
SO	Sequence 1260 BP; 407 A; 293 C; 258 G; 302 T; 0 other:		
 Query Match            62.5%; Score 1248.8; DB 20; Length 1260; Best Local Similarity   99.4%; Pred. No. 0; Matches 1253; Conservative   0; Mismatches     7; Indels       0; Gaps       0;			
OY	575	GAAACGATATGAGAAAGCATGTCATATACCAACTTACTGATGAGCATACAAAAGTCTCAAATCA	634
DB	1	gaacgttbtgaaaagcaqctcatccacaaattactgtagtgaactcaaatcacaagtccaatca	60

QY	635	GGAGTATTATTC	CCCGGCTTAC	TGAGTAATTT	TACTCTGA	AGGAAGT	CAATCAAT	GAAGTTGA	694
Dd	61	gcaagtaata	taccctgctt	aaaygaa	taattaccca	aggaagta	ctaaacaa	gtgaat	120
QY	695	GAAGAACA	AGGCTTAA	CCCTACT	CAACAAC	GGGANT	GGATGCT	TCTACACAAT	754
Dd	121	aaagaaca	gagctta	aaocctt	aaocctt	aaocctt	aaocctt	aaocctt	180
QY	755	GAAGCTCA	AAAAACAT	TCGTGT	GGGATAT	TATTA	CAATAC	AGAGAT	814
Dd	181	gaagctca	aaaaac	atcctt	ggttga	gaatatt	aaataca	agaatla	240
QY	815	GATGATTT	GCAAAGTC	GGCTTAC	CACTTGT	GATGTTT	CTAACG	TAAATCA	874
Dd	241	gatgatatt	gcgaatgc	atgcctt	caactt	ggttga	gaatatt	aaataca	300
QY	875	CTAGAGCA	AGCCCAT	TACAGTA	CAATAT	GTTCCT	TGGGAAT	TACCA	934
Dd	301	ctagagaca	gcgcctt	aaagta	gaatatt	cttccct	ggaatla	ccaaga	360
QY	935	AACCGCA	CTGGGANT	CAACAT	TAATAA	CCGATAC	AGAGAT	ATCTCT	994
Dd	361	aacgcgcac	tgttga	taacta	tgaaac	gcgata	caagata	tgcttc	420
QY	995	GGTGCT	CTACGAT	TCACTG	CTACTAT	GTGTTAC	AGATG	AGCCCT	1054
Dd	421	ggtgtctag	atgcctt	aaatgc	atgcctt	aaatgc	atgcctt	aaatgc	480
QY	1055	AATACTCT	GTATT	TAACTG	GGATAG	GGGGCT	ACTTTGG	CAACAT	1114
Dd	481	aatactct	gttctt	aaatgc	atgcctt	aaatgc	atgcctt	aaatgc	540
QY	1115	CTGCACA	TATG	GCACAAC	GTGCC	CGGTG	GAACAT	CTAAC	1174
Dd	541	ctgcaca	atctgc	gaagtc	gcgcgc	gcgcgc	gcgcgc	gcgcgc	600
QY	1175	GCCAAG	ACTTTCT	CTAAAT	TGCTT	AGATG	CACTAC	CCAGT	1234
Dd	601	gcacaag	acttctt	aaatgc	atgcctt	aaatgc	atgcctt	aaatgc	660
QY	1235	ATTTC	AATTA	CAAC	ACAC	CGAAT	CAAC	AAAAAT	1294
Dd	661	atttcaag	tataca	aaaccc	gaatgc	atgcata	aaataa	tggaag	720
QY	1295	GCTGCT	TACGCT	GGCTT	TGCAAA	TGTG	GAAC	TACTAT	1354
Dd	721	gctgctt	acgctt	gcctt	gcaaa	tgctga	actt	aaataa	780
QY	1355	GTCGCT	TATG	TATG	GGAG	TGAAAA	AGATCT	CTAAT	1414
Dd	781	gtcgcctt	atgcctt	aaatgc	atgcctt	aaatgc	atgcctt	aaatgc	840
QY	1415	GAAGA	CAGAC	CGCTT	ATATG	ATGAC	CGAGAT	ATAAAA	1474
Dd	841	gaacaaca	gcata	tatgat	gacgcata	gata	tgaaaa	acgctt	900
QY	1475	CGAAT	GGCTAT	CTTGG	TCTGCT	CCAG	GGCTG	GTAAAA	1534
Dd	901	cgaatg	ccata	cttctt	gcttgc	cttcc	ccagtc	tgtaaa	960
QY	1535	GACG	GAAG	AAAT	TGAAAC	CAAT	CAAG	ACCTTC	1594
Dd	961	gacgagga	aatatg	aaaccca	ccaaccca	ccctt	ccaat	tgtag	1020
QY	1595	GCTGG	CTATAC	GGCTT	AAATTT	CAATG	CGCTAT	TGAC	1654
Dd	1021	gctgctata	gcgtt	aaata	ttaacat	tgctt	gata	tgcaag	1080
QY	1655	CCAAT	TGAG	GCAT	TGGCCT	TAGG	CGCTG	CGCAAG	1714
Dd	1081	ccaat	tgagca	atctgc	cttgc	cttgc	cttgc	cttgc	1140

QY	1715	CAGTCTGAGGAAGCAATCCAGAAAGTTGGATATATCCAGAGGGGCTCTACAGAAATGGA	1774
DB	1141	CGTCTGGAAGGAAGCAATCCAGAGAGTCTGGAATATACCAAGAGGCTCTACAGAAATGGA	1200
QY	1775	GAAATTCGTATTTAAAAATGATGCTCTTCTACGTGGAACACCTGCTCCACAAACACC	1834
DB	1201	gaattcgtatttaaaatggtgctgtcttaagctggaactcaccctgtccacaacaacc	1260
RESULT 12			
ID	AAZ35951	standard; DNA; 1260 BP.	
XX	AAZ35951;		
AC	AAZ35951;		
XX	07-FEB-2000	(first entry)	
DT	07-FEB-2000	(first entry)	
DE	Streptococcus pneumoniae pbp1A TER isolate m)	nucleotide sequence.	
XX	Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;		
KW	transcriptase encoding region; TER; antibiotic resistance; diagnosis;		
KM	detection; identification; pneumococcal meningitis; ss.		
XX	Streptococcus pneumoniae.		
OS	Streptococcus pneumoniae.		
PN	ZA9807024-A.		
PN	28-APR-1999.		
XX	05-AUG-1998;	982A-0007024.	
PF	05-AUG-1998;	982A-0007024.	
XX	01-AUG-1997;	972A-0006886.	
PR	01-AUG-1997;	972A-0006886.	
XX	(SAME-) SOUTH AFRICAN INST MEDICAL RES.		
PA	(UYVI-) UNIV WITWATERSRAND.		
PA	(MEDI-) MEDICAL RES COUNCIL.		
XX	Klugman KP, Smith AM, Du Plessis M;		
PI	WPI; 1999-601770/51.		
XX	Polymerase chain reaction assays for detecting Streptococcus pneumonia		
PT	useful for the diagnosis of pneumococcal meningitis		
XX	Claim 11; Fig 4; 63pp; English.		
PS	A polymerase chain reaction (PCR) assays have been developed for		
XX	detecting an antibiotic resistant strain of Streptococcus pneumoniae		
CC	using primers based on the penicillin binding protein 2B (pbp2B) gene		
CC	and the pbp1A gene. The products and methods can be used for detecting		
CC	S. pneumoniae, particularly antibiotic-resistant strains and		
CC	used for simultaneously diagnosing pneumococcal meningitis and		
CC	identifying any antibiotic-resistant S. pneumoniae strains in a sample.		
CC	The methods can be used for detecting S. pneumoniae strains resistant		
CC	to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.		
CC	The assays can be adapted to detect other pathogens causing meningitis.		
CC	The assays can be used to detect an antibiotic resistant strain of		
CC	S. pneumoniae with a minimum inhibitory concentration (MIC) of		
CC	0.25-1 micro g/ml where the PCR products concentrated are a 1043 bp and		
CC	a 224 bp product. The present sequence represents a Streptococcus		
CC	pneumoniae pbp1A transcriptase encoding region (TER) isolate nucleotide		
CC	sequence from the present invention.		
XX	Sequence 1260 BP; 407 A; 293 C; 258 G; 302 T; 0 other;		
QY	Query Match	62.5%; Score 1248.8; DB 20; Length 1260;	
DB	Best Local Similarity	99.4%; Pred. No. 0;	
	Matches 1253; Conservative	0; Mismatches 7; Indels	0; Gaps
QY	575	GACAGTATGAGAAAGCACTCAATACCAATTCATGATGAGACTACAAAGTCTCAATCA	634
DB	1	gaacgtatgagaaagcaatcaatcaccaattacgtgagactaaagtcctcaatca	60

```
QY 635 GCAAGTAATTAACCTGGTTACATGATATTAATTAACCTCAAGAGAGTGCATCAATCAAGTTGAA 694
    |||
Db 61 gcaagtaataacctgctctacatgataataatactcaagaagatcaatcaatgaa 120
QY 695 GAAGAAACAGGCTATAAAGTCTACTGACAACTGGAGTGTCTACACAAATGTAAGACAA 754
    |||
Db 121 aagaacaagcgtataacccgtccacaacgtgagatgtctacacaatgtagaccaa 180
QY 755 GAAGCTCAAAAACATCTGTGGGATATTTACATACAGACAAATACGTTGGCTATCCAGAC 814
    |||
Db 181 gaagctcaaaaacatcgtggaatattacaatacagaagaatcgttgcctaccagac 240
QY 815 GATGAATTGCAATCGCTTACCATGTTGATGTTCTTAACGGTAAAGCATTTGCCAG 874
    |||
Db 241 gatgaattgcaaatgctcttaccatgttgaatgttcttaacggttaagatcattggccag 300
QY 875 CTAGAGACAGCCATCAGTACGTAATGTTCTTCGGAATTAACCAAGCAGTAAAGAAACA 934
    |||
Db 301 ctagagacagccatcagtaagtaatgttcccttcggaatcaacaaagcagtagaaca 360
QY 935 AACCGGACTGGGGATCACTATGAAACCGATACAGACTATGCTCTGCTTGGAGTAC 994
    |||
Db 361 aaccggactggggatcaactaagaaacgatacagactatgtctcctctgagttac 420
QY 995 GGTGTACAGATTCACCTGCTACTATATGTTTCAGGATGAGCCCTATTAACCTGCGGACA 1054
    |||
Db 421 ggtgtacagatccatcagctacatcgttcacagatgagccctatactaccctgggaca 480
QY 1055 AATACTCTGTTTATTAAGTGGATAGGGGCTACTTTGGCAACATCACTGCAATACGCC 1114
    |||
Db 481 aatactctgtttataactagtgataggggtacttcttgcaacatcaccttgcataagcc 540
QY 1115 CTGCAACATCGGAAACGTCCAGCCGTGAAACTCTAAACAAGGTGCACTCAACGCC 1174
    |||
Db 541 ctgcaacaatcggaaacgtccagccgtgaaactctaaacagtcgagactcaacgcg 600
QY 1175 GCCAAGCTTCTTAATGGTCTAGGAATGAGTACCTACCCAGTAATTCCTCAATAAGGCC 1234
    |||
Db 601 gccaaagcttcttaaatggctcggaaatcgactaacccaagatctcaactcaaaagcc 660
QY 1235 ATTTCAAGTACACAAACGATCAGACAAAAAATATGAGCAAGTATGAAAAAGATGCT 1294
    |||
Db 661 atttcaagtaacaacaacgagatcagacaaaaaatatgagcaagtagtgaaagatgct 720
QY 1295 GCTGCTTACGCTGCTTTGCAAAATGCTGCACTTACTATTAACCAATGTAATTCATTA 1354
    |||
Db 721 gctgcttacgctgctcttgcaaaatgtagaactataacaaatgatatccataaa 780
QY 1355 GTGCTGTTGTGATGGAGTGAATAAGTTCCTTAATGTGGAATTCGTGGCACTGAGTAA 1414
    |||
Db 781 gtcgtctttagtgatggagtgaaaaagatctctaaatgctggaactcgtgccatgaag 840
QY 1415 GAAACGACAGCTATATGATGACGACGATGATGAAAAACAGTCTGACTTGTGNAACGGA 1474
    |||
Db 841 gaaacgaacagccatataatgatacagacatgataaaccagtlcttgatgtgaactgga 900
QY 1475 CGAAATGCTATATGCTTGGCTGCTCCAGCGTGGTAAACAGAGACCTGTAACCTATACA 1534
    |||
Db 901 cgaatgctcatctgtcttgctccctcagcgtgtaaaacagaaaccttaactataca 960
QY 1535 GAGGAGGAATTTGAAAAACGATCAAGACTGCAATTTTACGACTGATGTAACCTTTT 1594
    |||
Db 961 gaggaggaatctgaaaaacacacaaagacctccaattttagacacctgacactattt 1020
QY 1595 GCTGGCTATACGGGTAAATATTCATGAGCTGTATGACAGGCTATTTCAACCGCTTGACA 1654
    |||
Db 1021 gctggctatacgcgtataataatcaatgctgtatgacagcgtatcttaacgcgttgaca 1080
QY 1655 CGACTGTAGGCAATGCGCTTACAGGTGCTGCGCAAAAGTTTACCGCTCTATGATGACCTAC 1714
    |||
Db 1081 ccactttaggaagcgtcttacgctcgtccaaagtattacgcgtctatgatgacctac 1140
```

```
QY 1715 CTGCTGAAGGAAGCAATCCAGAGATTGGAATATATACAGAGGGGCTCTACAGAAATGGA 1774
    |||
Db 1141 ctgctgaaaggaagcaatccagagatggaataatacagaaggggctctacagaagaa 1200
QY 1775 GAATTCGTAATTAATAAATGCTGCTGTTCTACGTTGGAAGTACCTGCTCCACAACACCC 1834
    |||
Db 1201 gaattcgatttaaaaaatgctgctgcttcaagcgtggaactcaactgctccacaacacc 1260

RESULT 13
AAZ35953
ID AAZ35953 standard; DNA; 1260 BP.
XX
AC AAZ35953;
XX
DT 07-FEB-2000 (first entry)
XX
XX Streptococcus pneumoniae pbp1A TER isolate o) nucleotide sequence.
DE
XX Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
KW transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KW detection; identification; pneumococcal meningitis; ss.
XX
OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.
PD
PD 28-APR-1999.
XX
PF 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 972A-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UWI-) UNIV WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX WPI; 1999-601770/51.
DR
XX
XX Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
PT
XX
XX Claim 11: Fig 4; 63pp; English.
PS
XX
XX A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1A gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
SQ Sequence 1260 BP; 405 A; 293 C; 260 G; 302 T; 0 other;

Query Match 62.5%; Score 1248.8; DB 20; Length 1260;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 575 GAACAGTATGAGAAGCAGTCAATACCAATTAATGATGAGTACCAAAAGTCTCAATCA 634
    |||
Db 1 gaacagtatgagaagcagtcataacacaaattactgtagtggtctacaagaagtcataatca 60
```

QY 635 GCAAGTAATTACCTGCTTACATGATGATTAATTACCTCAAGAGAGTCATCAATGAATTGAA 694  
|||||  
Db 61 gcaagtaattaccctgcttaccatgataatctaccctaaggaagtcatacaatgaa 120  
QY 695 GAAGAAACAGGCTATTAACCTACTACACACACGAGATGATCTACACAAATGAGACCA 754  
|||||  
Db 121 gaagaaacaggctataaccctaccacaactggaatgagtcctaccacaatgtagaacca 180  
QY 755 GAAGCTCAAAAACATCTGTGGATTTTACATACAGACAGATACGTTGGCTTACAGAC 814  
|||||  
Db 181 gaagctcaaaaaacatctgaggatalttaacaacagaacatgcttaccctaccagac 240  
QY 815 GATGAATTCGAAGTCGCTTCTACATGTTGATGTTTCTAACGCTAAAGTCATTTGCCAG 874  
|||||  
Db 241 gatgaattcgaaagtcgcttaccatctgtgatgtcttcaagctaaagcatcgccag 300  
QY 875 CTAGAGACAGCCATCAGTCAAGTATGTTTCTTGGGAATTAACCAAGCATAGAAACA 934  
|||||  
Db 301 cttagagcacgcacatcagtaagtaattctcttcggaalttaaccaagcagtagaaca 360  
QY 935 AACCGGACGTGGGATCACTATGAAACGATCAGACATATGCTGCTGCTGGAGTAC 994  
|||||  
Db 361 aaccggaactggggatcaactatgaaaccgatcaagactatgcttcgcttgaggatc 420  
QY 995 GGTGTCTACGATTCAGCTGCTACTATGTTTACAGATGAGCCCTATTAACCTGGGACA 1054  
|||||  
Db 421 ggtgtctacgatctcaactgctactatcgcttcaacgagccatactaccctgggaca 480  
QY 1055 AATACCTCCTGTTTATTAACAGTGGATAGGGCTACTTTGGCACTACCTTTGCAATACGC 1114  
|||||  
Db 481 aataccccctgtataacacggatagggatgactcttggacaacatccactgtcaatagcc 540  
QY 1115 CTGCAACAATCGGAAGAGCTCCAGCGGTGGAACACTTAAACAGGTCGAGCTCAACCGC 1174  
|||||  
Db 541 ctgcaacaatcgggaacgctccagccgtggaaaccttaacaagctcggaactcaaccgcg 600  
QY 1175 GCCAAGACTTTTCTTAATGTTAGTAATGACGTACCCCAAGTATTCATCTCAATAGCC 1234  
|||||  
Db 601 gccaaagacttccctaattgctcgaatcgactaccgaatcttcaactacctaagtcc 660  
QY 1235 AATTGACATACACACACCAATCAGACCAAAAATATGAGCAAGTATGTAAGAAAGATGCC 1294  
|||||  
Db 661 attcaagtaacacacaacgcaatcagaacaaaataatgagcaagtgtgaaagaagtgct 720  
QY 1295 GCTGCTTACGCTGCTTCTCAATAGTGGACCTTACTATAACCAATGATATCCATAA 1354  
|||||  
Db 721 gctgcttaagctgcttgcataatggttgaaacttactataaaccatgatatccataaa 780  
QY 1355 GTGCTCTTATGATGAGGAGTGAAGAAAGATCTCTTAATGTGGAACCTGTCGCATGAAG 1414  
|||||  
Db 781 gtcgctcttagtgatggagtgaaaagagttctcttaactgctgaactgctgcatgaaa 840  
QY 1415 GAAAGACAGCCTATATGATGACCCGACATGATGAAGAAAGCTCTTGACTTATGGAAGTGA 1474  
|||||  
Db 841 gaaagacacgacctatattgtgacccagacatgataaagagctcttgattatgaaactgga 900  
QY 1475 GCAAAATGCTATCTTGGCTGGCTCCCTGAGGCTGTAAGAAAGCAAGCACTCTAATCTTAA 1534  
|||||  
Db 901 gcaaatgctactctgtctggtccctcagcggtaaaaacaggaacccctatactaca 960  
QY 1535 GACGAGGAATTAAGAAACACATCAAGACCTCTCAATTTGTAGCACCTGATGAATCTT 1594  
|||||  
Db 961 gacgaggaattgaaaacacacatcaagacctctcaattgtagcactctgtagaactatt 1020  
QY 1595 GCTGCTATACCGCTTAATATTTCAATGCTGTATGACAGGCTATTTTACCGCTGACA 1654  
|||||  
Db 1021 gctgctataacgcgtaaaatactcaatggtgatgacaggtatcttaaccgctgaca 1080  
QY 1655 CCACCTGTATGGAATGGCTTACGGCTGCGCAAAAGTTTACCGCTCTGATGAGCTAC 1714  
|||||  
Db 1081 ccactctgtaggaatgagccctacggtcgctgccaagaagttaccgctctatgtatgacctac 1140

QY 1715 CTGCTGAAAGAACCAATCCAGAGATTGGAATATACCAAGGCGCTTACAGAAATGCA 1774  
|||||  
Db 1141 ctgctgaaaggaagcaatccacagagattggaatataccagaaggcgctctacagaatgga 1200  
QY 1775 GAATTCGATTATTAATAATGCTGCTGTTCTACTGGAACCTACTGCTCACAACAACCC 1834  
|||||  
Db 1201 gaatcgtattataaataatgctgctgcttactcgtgagctcacctcctccacaacacc 1260

RESULT 14  
AAH02067  
ID AAH02067 standard; DNA: 1249 BP.  
XX  
AC AAH02067;  
XX  
DT 24-JUL-2001 (first entry)  
XX  
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2060.  
XX  
KW Species specific; genus specific; family specific; probe; detection;  
KW Identification; algal; archaeal; bacterial; fungal; parasitical;  
KW microorganism; diagnosis; translation elongation factor Tu; toxin;  
KW translation elongation factor G; RecA recombinase; resistance;  
KW catalytic subunit of proton-translocating ATPase; antimicrobial;  
KW vaccine; primer; ds.  
XX  
OS Streptococcus pneumoniae.  
XX  
PN WO200123604-A2.  
XX  
PD 05-APR-2001.  
XX  
PF 28-SEP-2000; 2000WO-CA01150.  
XX  
PR 28-SEP-1999; 99CA-2283458.  
PR 19-MAY-2000; 2000CA-2307010.  
XX  
XX  
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.  
PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;  
PI Picard FJ, Roy PH;  
XX  
DR WPI: 2001-245006/25.  
XX  
XX  
PT Nucleic acid sequences are used to generate universal probes and  
PT primers which can be used to identify and detect the presence of algal,  
PT archaeal, bacterial, fungal and parasitical species in a test sample -  
XX  
PS Claim 27; Page 1473; 1580bp; English.

The present invention describes a method for generating a repertory of nucleic acids of tuf, fus, atp and/or recA genes from which probes and/or primers are derived. The method comprises amplifying the nucleic acids of determined algal, archaeal, bacterial, fungal and parasitical species with a combination of defined primer pairs. The method can be used for producing probes and/or primers for detecting one or more related microorganisms e.g. algae, archaea, bacteria, fungi and parasites, for universal detection and for specific and ubiquitous detection and identification of an algal, archaeal, bacterial, fungal and parasitical species, genus, family and group. A nucleic acid (1) obtained using the method of the invention can be used for the universal detection of any bacterium, fungus or parasite in a sample and for the detection of at least one antimicrobial agent resistance gene or at least one toxin gene. hexa nucleic acids are used for the specific and ubiquitous detection and for identification of Streptococcus pneumoniae. (1) can be used to design a therapeutic agent which is effective against microorganisms. Microbial species or genus or family or phylum or group which can be detected include Abiotrophia adiacens, Bordetella sp., Corynebacterium sp., Enterobacteriaceae group, Escherichia coli, Mycobacteriaceae family, Pseudomonads group, Streptococcus sp., Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests provides faster results than substrate specificity tests as results can be determined in an hour and improved accuracy is also achieved.

AAH00010 to AAH002304 represent nucleotide sequences and primers/probes which are given in the exemplification of the present invention.  
 CC Best Local Similarity 99.9%; Pred. No. 0;  
 XX Matches 1248; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 SQ Sequence 1249 BP; 402 A; 302 C; 246 G; 299 T; 0 other;

Query Match 62.4%; Score 1247.4; DB 22; Length 1249;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 XX Matches 1248; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 SQ Sequence 1249 BP; 402 A; 302 C; 246 G; 299 T; 0 other;

720 CAACCTGGGATGATGCTACACAAATGTAGACCAGAGAGCTCAAAAACATCTGTGGAGTA 779  
 1 caactgysatgagctctacacaaatgtagaccagaagctcaaaaacatctgtggagata 60  
 780 TTACCAATGACAGCAATGCTTGGCTATCCAGAGATGAATTTGCAAGTCGCTTACCA 839  
 61 ttacaaatacagagaaatgctgctacccagagaaatgtagaagctgcttccacaa 120  
 840 TTGTTGATGTTCTTAACGGTAAAGTCATTGCGCAGCTAGAGACGCCATCACTCAAGTA 899  
 121 ttgttgatgttcttaacggtaaaagtcatgtccagctagagagacgcatcagtcagta 180  
 900 ATGTTTCCTTGGGATTAACCAAGCAAGTAAACAAACCGGACCTGGGATCACTATGA 959  
 181 atgttcccttcggaatlaaccaaagcagtagaaacaaacgcgactgggagatcaactga 240  
 960 AACCGATCAGACATGATGCTGCTGCTTGGAGTACGCTTACGATTTCAACGCTACTA 1019  
 241 aacgcagacacagactatgctcctgctgagtaagtgctacgaltcaacagctactaa 300  
 1020 TCGTTACAGATGAGCCCTTAATTAACCTCGGACAAATATCTCCTGTTTAACTGGGATA 1079  
 301 tctgtcacagatgagcccttaactaactccctgggacaataatctctgtttataactggata 360  
 1080 GGGGCTACTTGTGGCAATCATTGCTGCAATGAGCCCTGCAACAACTGCGAAACGTCCTCA 1139  
 361 ggggctacttggcaaatcattgcaatgacgctcctgcaacatcgcaaacgctccag 420  
 1140 CGGTGGAACCTGTAACAGAGTGGGCTGACGCGGCAAGCTTCTGTAATAGTGTCTG 1199  
 421 ccgtggaacactttaaacaagctgagactcaacgcgcaagacttccataatgctcag 480  
 1200 GAATGACTACCAAGTATTTCACTACTCAATAGCCATTTCAAGTAACACACGGAATCAG 1259  
 481 gaatgactacccaagatattcaactactcaaatgcaatccttcaagtaacaaacggaatcag 540  
 1260 ACAAATAATATGAGCAAGTAGTGAAGAGATGCTGCTTACGCTGCTTGAATG 1319  
 541 acaaaaataatgtagcaagtagtgaagaagatgctgcttgaagctgcttgccttgcgaatg 600  
 1320 GTGGAACCTACTATAACCAATGTATATCCATTAAGTCTTCTTATGATGGAGTGAAT 1379  
 601 gtggaacactactataaacaatgataatcaatgcttcttgaatgagtgagaa 660  
 1380 AAGAGTCTCTATGTCGGAAGCTGCGCATGGAAGGAAGACAGAGCTTATGATGATACCG 1439  
 661 aagagctctctaatgctggaactgctgcaatgaaagaaacgacagcctataatgataacg 720  
 1440 ACATGATGAAAAAGTCTTGAATGATGAACTGGAGAAATGCTTCTTGTGGCTCC 1499  
 721 acatgctgaaaaagcttctgactatgtaactggaactggaactgcttcttgcctcc 780  
 1500 CTCAGGCTGTGTAACAGACACCTCTAATATACAGACAGAAATTTGAAACACATCA 1559  
 781 ctcaggctgtgtaaaacagaaactctaaactatacagagaaatgtaaaacacataca 840  
 1560 AGACCTCTCAATTTGTAGACCTGATGATGATTTGCTGGCTATACGCTAAATATTCA 1619  
 841 agacctctcaattttagaactgataactgataacttgcgtcactaagcgtataatcaatca 900  
 1620 TGGCTGTATGACAGAGCTATTTCAACCGTGTACACACCACTTGTAGCAATGGCTTACGG 1679  
 901 tggctgtatgacagagctatcttaaccgctgacacacttgaagcaatgagcttaacg 960

1680 TCGCTCCAAAGTTTACCGCTGATGATGACCTACCTGCTGTAAGGAGCAATCCAGAA 1739  
 961 tgcctccaaagtcttaacgctctgataatgacactctgctgaaagaaatccagaag 1020  
 1740 ATTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTATTAATAATGCTCTC 1799  
 1021 attggaatatcagaggggctctacagaatgagaaatcgtatcttataaataatgctc 1080  
 1800 GTTCTACGTGGAATCACTGCTCCACAAACACCCCATCACTGTAAGTTCAAGCTCAT 1859  
 1081 gttctacgtgagctcactgctccacaacaccccatcaactgaaatgtaagctcat 1140  
 1860 CATCAGATGATTCACATTCACAGCTAGCTAGCTCAACCACTCCAAAGCAAAATATAGTACGA 1919  
 1141 catcagatagttcaacttcaacgctcagctcaacacccatcaacagacaaataatagta 1200  
 1920 CTACCAATCTTAACAATATACGCAACAAATCAAAATCAACCCCTGATCA 1968  
 1201 ctaccatctctaataataatagcgaacatcaatacaacccctgata 1249

RESULT 15  
 AA235939  
 ID AA235939 standard; DNA; 1260 BP.  
 XX  
 AC AA235939;  
 XX  
 DT 07-FEB-2000 (first entry)  
 XX  
 DE Streptococcus pneumoniae pbp1A TER isolate a) nucleotide sequence.  
 XX  
 KW Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;  
 KM transpeptidase encoding region; TER; antibiotic resistance; diagnosis;  
 KW detection; identification; pneumococcal meningitis; ss.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 FH Location/Qualifiers  
 FT 1..1260  
 FT CDS  
 FT /\*tag= a  
 FT /product= "pbp1A TER"  
 FT /note= "no stop codon given"  
 XX  
 PN ZA9807024-A.  
 XX  
 PD 28-APR-1999.  
 XX  
 PF 05-AUG-1998; 982A-0007024.  
 XX  
 PR 01-AUG-1997; 972A-0006886.  
 XX  
 PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.  
 PA (UYWI-) UNIV WITWATERSRAND.  
 PA (MED-) MEDICAL RES COUNCIL.  
 XX  
 PI Klugman KP, Smith AM, Du Plessis M;  
 XX  
 XX WPI: 1999-601770/51.  
 DR P-PSDB; AAY56106.  
 DR  
 XX  
 PT Polymerase chain reaction assays for detecting Streptococcus pneumonia  
 PT useful for the diagnosis of pneumococcal meningitis  
 XX  
 PS Claim 11: Fig 4; 63bp; English.  
 XX  
 CC A polymerase chain reaction (PCR) assays have been developed for  
 CC detecting an antibiotic resistant strain of Streptococcus pneumoniae  
 CC using primers based on the penicillin binding protein 2B (pbp2B) gene  
 CC and the pbp1A gene. The products and methods can be used for detecting  
 CC S. pneumoniae, particularly antibiotic-resistant strains. They can be  
 CC used for simultaneously diagnosing pneumococcal meningitis and  
 CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.

